Paper 95

By: Trial Section Merits Panel
Board of Patent Appeals and Interferences
Mail Stop Interference
P.O. Box 1450
Alexandria, VA 22313-1450
Tel: 571-272-9797
Fax: 571-273-0042

Filed: March 9, 2007

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE BOARD OF PATENT APPEALS
AND INTERFERENCES
(Administrative Patent Judge Richard E. Schafer)

Human Genome Sciences, Inc.,
Junior Party
(Application 09/042,583-IFW
Inventors: JIAN NI, REINER L. GENTZ,
GUO-LIANG YU and CRAIG A. ROSEN),

v.

Immunex Corp.,
Senior Party
(Patent 6,569,642
Inventors: CHARLES RAUCH and HENNING WALCZAK).

Patent Interference No. 105,380 (RES)

Before: SCHAFER, HANLON and SPIEGEL, Administrative Patent Judges.

SPIEGEL, Administrative Patent Judge.

DECISION - MOTIONS - Bd.R. 125(a)

I. Introduction

This is a decision on the motions remaining in interference 105,380.

Junior party Ni has filed four motions, one of which has been previously decided.[1] Senior party Rauch has filed five motions.

Ni substantive motion 1 to substitute proposed count 2 for Count 1 is **denied**. Ni substantive motion 2 to be accorded priority benefit of the 17 March 1997 and 29 July 1997 filing dates of its U.S. provisional applications 60/040,846 and 60/054,021, respectively, is **dismissed** as moot as to Ni's proposed count 2, **granted** as to application 60/054,021 for Count 1, and **denied** as to application 60/040,846. Ni miscellaneous motion 4 to exclude certain evidence is **denied**.

Rauch substantive motion 1 to be accorded priority benefit of the 4 June 1997 and 28 March 1997 filing dates of its U.S. applications 08/869,852 and 08/829,536, respectively, as to Count 1 is **granted**. Rauch substantive motion 2 to designate Ni claims 321, 322, 324 and 477 of the '583 application as corresponding to Count 1 is **denied**. Rauch substantive motion 3 for judgment that Ni claims 287, 289-299, 351-361, 389-403, 431, 432, 434-442, 446, 448-458, 507-517, 581 and 623-632 are unpatentable under 35 U.S.C. §§ 102(a) and/or (e) over any one of U.S. Patents 6,642,358, 6,072,047 and 6,569,642 and WO 98/35986 is **granted to the extent** that the claims 287, 289-299, 351-361, 389-403, 431, 432, 434-442, 446, 448-458, 507-517, 581 and 623-632 are unpatentable under § 102(e) over U.S. Patent 6,072,047. Rauch responsive motion 4 for priority benefit as to Ni's proposed count 2 is **dismissed** as moot.

---

[1] Ni substantive motion 3 (Paper 30) to hold each of Rauch's involved claims unpatentable under 35 U.S.C. § 102(e) as anticipated by U.S. Patent 6,872,568 has already been denied (Paper 46).

Rauch miscellaneous motion 5 to exclude certain evidence is **dismissed** as moot.

II. Findings of Fact (FF)

The following findings of fact are supported by a preponderance of the evidence.

1. The junior party is Jian NI, Reiner L. GENTZ, Guo-Liang YU and Craig A. ROSEN ("**Ni**").
2. Ni is involved in the interference on the basis of application 09/042,583 ("the "583 application," NX 2024), filed 17 March 1998.
3. Ni's real party-in-interest is Human Genome Sciences, Inc. ("HGS").
4. The senior party is Charles RAUCH and Henning WALCZAK ("**Rauch**").
5. Rauch is involved in the interference on the basis of U.S. Patent 6,569,642 ("the '642 patent," RX 1008), issued 27 May 2003, based on application 09/536,201 ("the '201 application"), filed 27 March 2000.
6. The '201 application has been accorded benefit for the purpose of priority of the 26 June 1997 filing date of application 08/883,036, which issued 6 June 2000 as U.S. Patent 6,072,047 ("the '047 patent," RX 1048).
7. Rauch's real party-in-interest is Immunex Corp. ("Immunex").
8. The subject matter of the interference is defined by one count.
9. Count 1 is "Claim 5 of U.S. Patent 6,569,642" (Paper 1, p. 3).
10. Claim 5 of the '642 patent, written in independent form, reads:

> An isolated TRAIL-R DNA, wherein said DNA encodes a polypeptide comprising an amino acid sequence that is at least 90% identical to the amino acid sequence present in SEQ ID NO:2, wherein said polypeptide binds TRAIL.

11. According to the '642 patent, SEQ ID NO:2 is the 440 amino acid sequence of a full length human receptor protein (including the N-terminal signal peptide), "TRAIL-R," encoded by the DNA of SEQ ID NO:1 (RX 1008, c. 1, l. 66 - c. 2, l. 2 and c. 22, ll. 12-14).

12. The claims of the parties are:

Ni	287-299, 319-322, 324, 326-339, 351-361, 389-403, 431-432, 434-442, 446-458, 476-491, 507-517, 553-596, 598-607 and 623-632
Rauch	1-52

13. The claims of the parties which correspond to Count 1 are:

Ni	287, 289-299, 351-361, 389-403, 431-432, 434-442, 446, 448-458, 507-517, 580-596, 598-607 and 623-632
Rauch	1-5, 7-13, 15-16, 18, 22, 25-26, 28-29, 31, 35, 38-42, 44, 48 and 51

14. The claims of the parties which do not correspond to Count 1, and therefore are not part of this interference, are:

Ni	288, 319-322, 324, 326-339, 447, 476-491 and 553-579
Rauch	6, 14, 17, 19-21, 23-24, 27, 30, 32-34, 36-37, 43, 45-47, 49-50 and 52

Other findings of fact follow below.

III. Ni Substantive Motion 1

Pursuant to 37 CFR § 41.121(a)(1)(i), Ni moves to redefine the scope of the interference by substituting proposed count 2 for current Count 1 (Paper 28). Rauch opposes (Paper 52); Ni replies (Paper 58).

15. Ni's proposed count 2 reads (Paper 28, p. 1, ¶ 1):

> An isolated TRAIL-R DNA, wherein said DNA encodes a polypeptide comprising an amino acid sequence that is at least 90% identical to the amino

> acid sequence presented in SEQ ID NO:2, wherein said polypeptide binds TRAIL or induces apoptosis.

16. According to Ni, its proposed count 2 simply incorporates Rauch claims 4 and 5, as does the current count, and adds the language "or induces apoptosis" (id.).

It is our understanding that the source of SEQ ID NO:2 in Ni's proposed count 2 is the involved '642 patent of Rauch. With this understanding, we now address Ni motion 1.

17. Ni argues that the abilities to bind TRAIL and to induce apoptosis are inherent properties of the polypeptide encoded by the DNA of Count 1, although only the former is expressly recited in the count (id., p. 8, ¶ 1).

A party seeking to change the count in an interference must demonstrate a genuine need to change the count. As stated in Louis v. Okada, 59 USPQ2d 1073, 1076 (Bd. Pat. App. & Int. 2001),

> [a]t a minimum, ... a preliminary motion to broaden out the count on the basis that a party's best or earliest proofs are outside of the current count (1) should make a proffer of the party's best proofs, (2) show that such best proofs indeed lie outside of the scope of the current count, and (3) further show that the proposed new count is not excessively broad with respect to what the party needs for its best proofs.

Ni seeks to change the count by adding the limitation "or induces apoptosis" as an alternative to the limitation "binds TRAIL" (FF 15). Ni seeks to change the current count because its best proofs do not explicitly recite that the TRAIL-R DNA of the count encodes a polypeptide that binds TRAIL (FF 17). However, the fact that Ni's "best proofs" do not explicitly recite the language of

the count does not alone establish that those proofs are not directed to "subject matter" defined by the count. "The invention is not the language of the count but the subject matter thereby defined." Silvestri v. Grant, 496 F.2d 593, 598, 181 USPQ 706, 709 (CCPA 1974). In appropriate circumstances, express limitations of the count may be shown to be inherent in the proofs. Id. ("In reaching this conclusion, we do not disregard the fact that the count also requires that the ampicillin possesses greater storage-stability than hydrated ampicillin and have a molecular weight of about 349. However, we regard these as inherent properties of Form II ampicillin which add nothing to the count definition beyond that determined by the [other limitations]."). [2] The limitation said not to be disclosed by Ni's best proofs, i.e., the ability to bind TRAIL, may be shown to be an inherent property of the polypeptide encoded by the DNA of the count. In fact, Ni argues that both the ability to bind TRAIL and the ability to induce apoptosis are both inherent properties of the polypeptides encoded by the DNA of the count:

> The ability to bind TRAIL is an expressly recited property of the polypeptides encoded by the DNA and it is an inherent property of the polypeptide of SEQ ID NO:2. Similarly, the ability of the polypeptide of SEQ ID NO:2 to induce apoptosis is also an inherent property of the polypeptide of SEQ ID NO:2.

[Paper 28, p. 8 (citation to material facts omitted).] Additionally, Ni has not asserted that there are any DNA molecules which would express polypeptides

[2] In Silvestri, the count was directed to a new crystalline form of ampicillin which was "substantially free of water in the chemically bound state" and had a molecular weight of about 349, a particular infrared ("IR") spectrograph and improved storage stability vis-à-vis the previously known form of ampicillin. Id., 496 F.2d at 595-96, 181 USPQ at 709-710. The court held that it was sufficient to possess the claimed compound and to characterize it by its water content and IR spectrograph, without demonstrating knowledge of the ampicillin's molecular weight because the molecular weight "add[s] nothing to the count beyond that determined by the water content and infrared spectrograph." Id., 496 F.2d at 599, 181 USPQ at 709.

meeting the amino acid sequence requirement of the count which would induce apoptosis, but not bind TRAIL. Consequently, adding the phrase "or induces apoptosis" to Count 1 has not been shown to be necessary to encompass Ni's best proofs. Furthermore, changing the scope of the count would leave Ni in essentially the same position it is in now of having to prove an inherent property of the receptor polypeptide encoded by the DNA of the count (FF 17). Hence, Ni has failed to demonstrate that its best proofs are outside the scope of the current count and, therefore, that there is a genuine need to change the count.

Based on the foregoing, Ni substantive motion 1 is **denied**.

IV. Rauch Responsive Motion 4

Pursuant to 37 CFR § 41.121(a)(2), Rauch moves to be accorded benefit for the purpose of priority of the 26 June 1997, 4 June 1997 and 28 March 1997 filing dates of U.S. applications 08/883,036, 08/869,852 and 08/829,536, respectively, as to Ni proposed count 2 (Paper 44). Rauch responsive motion 4 is contingent upon the grant of Ni substantive motion 1 to substitute Ni proposed count 2 for current Count 1. Since the contingency has not occurred, Rauch responsive motion 4 is **dismissed** as moot.

V. Rauch Substantive Motion 1

Pursuant to 37 CFR § 41.121(a)(1)(ii), Rauch moves to be accorded benefit for the purpose of priority of the 4 June 1997 and 28 March 1997 filing dates of U.S. applications 08/869,852 ("the '852 application," RX 1017) and 08/829,536 ("the '536 application," RX 1016), respectively, as to Count 1 (Paper 33). Ni opposes (Paper 48); Rauch replies (Paper 60).

The '201 application from which Rauch's involved '642 patent issued has already been accorded benefit of the 26 June 1997 filing date of Rauch's earlier filed '036 application (FF 6).

18. The '201 application is a continuation of the '036 application.

19. The disclosures of the '201 and '036 applications are substantially identical.

20. The '036 application is a continuation-in-part of the '852 application, which is a continuation-in-part of the '536 application.

21. The '852 application was filed 4 June 1997 (RX 1017, p. 55).

22. The '536 application was filed 28 March 1997 (RX 1016, p. 32).

To be accorded benefit for the purpose of priority in an interference proceeding "means Board recognition that a patent application provides a proper constructive reduction to practice under 35 U.S.C. 102(g)(1)." 37 CFR § 41.201. A constructive reduction to practice "means a described and enabled anticipation under 35 U.S.C. 102(g)(1) in a patent application of the subject matter of a count." Id. Benefit for the purpose of priority focuses on the subject matter of a count and only requires a constructive reduction to practice of a single embodiment within the scope of the count. Falkner v. Inglis, 463 F.3d 1376, 1379, 79 USPQ2d 1001, 1004 (Fed. Cir. 2006); Hunt v. Treppschuh, 523 F.2d 1386, 1389, 187 USPQ 426, 429 (CCPA 1975).[3]

The subject matter of Count 1 involves isolated DNA encoding a polypeptide having an amino acid sequence which (a) is at least 90% identical to SEQ ID

[3] In contrast, benefit for the purpose of 35 U.S.C. § 120 and related statutes focuses on the subject matter of the claim and requires the application for which benefit is sought to describe and enable the entire scope of the claim.

NO:2 of the '642 patent and (b) binds TRAIL (FF 10). SEQ ID NO:2 of the '642 patent is the amino acid sequence of a full-length human receptor protein called TRAIL-R and is encoded by the DNA of SEQ ID NO:1 in the '642 patent (FF 11).

23. It is undisputed that the TRAIL-R protein described in SEQ ID NO:2 of the involved '642 patent is the 440 amino acid isoform of a receptor protein alternatively referred to in the literature as TR-2, DR5, Apo-2, TRICK2 and KILLER (see Paper 48, p. B-1 where Ni admits Rauch Statement of Material Facts ("SMFs") 1 and 6 as set forth in Paper 33, p. 10).[4]

24. According to the '642 specification, TRAIL or "TNF-related apoptosis-inducing ligand" is a member of the tumor necrosis factor ("TNF") family of ligands and TRAIL-R binds TRAIL (RX 1008, c. 1, ll. 18-20 and 60-61).

25. Further according to the '642 specification, "[c]ertain uses of TRAIL-R flow from this ability to bind TRAIL, ...TRAIL-R finds use in inhibiting biological activities of TRAIL, or in purifying TRAIL by affinity chromatography, for example" (id., c. 1, ll. 61-65; these and additional uses are set forth at c. 15, l. 41 - c. 20, l. 18).

26. Example 6 in the '642 specification (c. 27, l. 26 - c. 28, l. 24) is said to demonstrate the ability of full length human TRAIL-R to bind TRAIL.

27. It was known at the time both the earlier '852 and '536 Rauch applications were filed that TRAIL was capable of inducing apoptosis (RX

[4] An isoform is one of the several forms in which a protein may exist in various tissues.

1026;[5] see also Paper 48, p. B-1 where Ni admits Rauch SMF 12 as set forth in Paper 33, p. 11).

28. SEQ ID NO:1 of the '852 application is said to present a DNA sequence encoding a human TRAIL receptor protein (TRAIL-R) having the amino acid sequence set forth in SEQ ID NO:2 of the '852 application (RX 1017, p. 13-15).

29. The '852 specification explicitly states that "[t]he present invention provides isolated nucleic acids useful in the production of TRAIL-R polypeptides, ... Such nucleic acids include, but are not limited to, the human TRAIL-R DNA of SEQ ID NO:1." (RX 1017, p. 27, ll. 26-29).

30. It is undisputed that the amino acid sequence described in SEQ ID NO:2 of the '852 application is identical to amino acid sequence SEQ ID NO:2 of the '642 patent (see Paper 48, B-1 where Ni admits Rauch SMF 10 as set forth in Paper 33, p. 11).

31. According to the '852 specification, TRAIL-R binds to the cytokine TRAIL and "[c]ertain uses of TRAIL-R flow from this ability to bind TRAIL, ... TRAIL-R finds use in inhibiting biological activities of TRAIL, or in purifying TRAIL by affinity chromatography, for example" (RX 1017, p. 2, ll. 8-12; these and additional uses of TRAIL-R are set forth at p. 20, l. 15 - p. 25, l. 14).

32. Example 6 in the '852 specification (p. 35, l. 4 - p. 36, l. 13) is said to demonstrate the ability of full length human TRAIL-R to bind TRAIL.

---

[5] Wiley et al., "Identification and Characterization of a New Member of the TNF Family that Induces Apoptosis," Immunity, Vol. 3, pp. 673-682 (December 1995) (RX 1026).

33. Thus, the '852 application describes an embodiment within the scope of Count 1, i.e., a DNA sequence encoding a human TRAIL receptor protein (TRAIL-R) having an amino acid sequence identical to the amino acid sequence of SEQ ID NO:2 of the '642 patent (FFs 28-30) and which binds TRAIL (FFs 31-32).

34. Ni does not dispute Rauch's claim to benefit for the purpose of priority of the filing date of its '852 application (Paper 48).

35. Figure 2 of the '536 application is said to present a DNA sequence encoding a human TRAIL receptor protein (TRAIL-R) having the amino acid sequence set forth in Figure 3 of the '536 application (RX 1016, p. 2, ll. 1-5).

36. The '536 specification explicitly states that "TRAIL-R DNA may be used to prepare TRAIL-R polypeptides encoded by the DNA" (RX 1016, p. 6, ll. 7-10).

37. It is undisputed that the nucleic acid and amino acid sequences of the DNA and encoded TRAIL-R protein described in Figures 2 and 3 of the '536 application are identical to the nucleic acid and amino acid sequences described in SEQ ID NOs: 1 and 2, respectively, of the '642 patent (see Paper 48, p. B-1 where Ni admits Rauch SMFs 13 and 14 as set forth in Paper 33, pp. 11-12).

38. According to the '536 specification, TRAIL-R binds to the cytokine TRAIL and "[c]ertain uses of TRAIL-R flow from this ability to bind TRAIL, ... TRAIL-R finds use in inhibiting biological activities of TRAIL, or in

purifying TRAIL by affinity chromatography, for example" (RX 1016, p. 2, ll. 8-12; these and additional uses are set forth at p. 13, l. 34 - p. 18, l. 26).

39. Thus, the '536 application describes an embodiment within the scope of Count 1, i.e., a DNA sequence encoding a human TRAIL receptor protein (TRAIL-R) having an amino acid sequence identical to the amino acid sequence of SEQ ID NO:2 of the '642 patent (FFs 35-37) and which binds TRAIL (FF 38).

40. Ni disputes Rauch's claim to benefit of the filing date of the '536 application, contending that the '536 application fails to show any utility for the DNA molecule and, therefore, fails the "how to use" of the enablement requirement (Paper 48, § C.2, pp. 5-6).

41. Ni relies on Rasmusson v. SmithKline Beecham Corp., 413 F.3d 1318, 75 USPQ2d 1297 (Fed. Cir. 2005) to support its conclusion that

> a party cannot establish that an earlier application constitutes a constructive reduction to practice without at the very least showing that the earlier application discloses a utility for an embodiment of the count. In other words, an essential element of a party's case for benefit of an earlier application is a demonstration that the earlier application satisfies the how-to-use prong of § 112, first paragraph, with respect to at least one embodiment of the count. [Paper 48, p. 5, ¶ 2.]

42. Specifically, Ni argues that "Rauch has neglected to assert, implicitly or explicitly, that the '536 application discloses any utility for a DNA molecule within the scope of the count" (Paper 48, p. 7, ¶ 3).

In essence, the only opposition raised by Ni is whether the '536 application discloses an adequate utility/enablement for a DNA embodiment within the scope of Count 1. First, count 1 explicitly describes a utility for a DNA embodiment within the scope of Count 1, i.e., it encodes a polypeptide having an amino acid sequence which is at least 90% identical to SEQ ID NO:2 of the '642 patent and binds TRAIL. Second, Rauch asserted this utility/enablement (Paper 33, pp. 5-6) and pointed to express descriptive support of an embodiment within the scope of Count 1 in the '536 application in Appendix C of its motion, i.e., "Figure 2 discloses a DNA sequence that encodes the polypeptide set forth in Figure 3" (Paper 33, p. 13, c. 2). Third, the '536 specification explicitly states that "TRAIL-R DNA may be used to prepare TRAIL-R polypeptides encoded by the DNA" (FF 36). Fourth, our finding that the '536 specification describes and enables an embodiment within the scope of Count 1 is not inconsistent with the holding in Rasmusson.

In Rasmusson, both parties had interfering claims directed to methods of treating prostate cancer comprising administering finasteride, a selective 5-α-reductase inhibitor. An interference was declared by the Board of Patent Appeals and Interferences ("the Board"). Rasmusson was involved in the interference on the basis of an application which claimed priority to eight earlier filed applications. SmithKline Beecham Corp. was involved in the interference on the basis of two patents and corresponding reissue applications. On appeal from the decision of the Board, the Federal Circuit affirmed the Board's holding that Rasmusson was not entitled to benefit for the purpose of priority of the filing

dates of the eight earlier filed applications. Citing In re Brana, 51 F.3d 1560, 34 USPQ2d 1436 (Fed. Cir. 1995), the court said "a specification disclosure which contains a teaching of the manner and process of making and using the invention . . . must be taken as in compliance with the enabling requirement of the first paragraph of § 112 unless there is reason to doubt the objective truth of the statements contained therein which must be relied on for enabling support" (Rasmusson, 413 F.3d at 1323, 75 USPQ2d at 1300, emphasis added). The court affirmed the Board's finding that one of ordinary skill in the art would not have believed that finasteride was effective in treating prostate cancer in light of the state of the art at the relevant time and because Rasmusson had failed to provide experimental proof demonstrating the effectiveness of the invention (id., 413 F.3d at 1324-25, 75 USPQ2d at 1301).

Here, the '536 specification explicitly states that "TRAIL-R DNA may be used to prepare TRAIL-R polypeptides encoded by the DNA" (FF 36). The '536 specification further describes certain uses of TRAIL-R based on its ability to bind TRAIL, e.g., using TRAIL-R to purify TRAIL by affinity chromatography (FF 38). Ni has not pointed to evidence of record which raises doubts as to the objective truth of these statements in the '536 specification, as was the case in Rasmusson. For example, Ni does not allege or provide evidence that one of ordinary skill in the art could not use a DNA to produce the protein encoded thereby or that a receptor protein that binds a ligand could not be used to purify the ligand using affinity chromatography at the time the '536 application was filed. Specifically, Ni has not shown that Rauch failed to satisfy its burden of proof with

respect to enablement regarding the DNA embodiment of the count set forth in Figure 2 of the '536 application, i.e., that the DNA of Figure 2 did not encode, and therefore was not useful to produce, the TRAIL-R polypeptide set forth in Figure 3. Moreover, Ni does not argue that the '536 application fails to disclose any utility for the TRAIL-R polypeptide set forth in its Figure 3. In short, Rauch described how to use a DNA within the scope of Count 1, i.e., the DNA of Figure 2 encodes a protein that binds TRAIL (Paper 33, pp. 5-6), and Ni has not asserted that encoding a protein that binds TRAIL is not a sufficient utility nor provided any basis to doubt the objective truth of express statements in the '536 specification that the DNA of Figure 2 is useful to produce the encoded TRAIL-R polypeptide.

Based on the foregoing, Rauch substantive motion 1 is **granted**.

VI. Ni Substantive Motion 2

Pursuant to 37 CFR § 41.121(a)(1)(ii), Ni moves to be accorded benefit for the purpose of priority of the 17 March 1997 and 29 July 1997 filing dates of its earlier provisional applications 60/040,846 ("the '846 application," NX 2042) and 60/054,021 ("the '021 application," NX 2056), respectively, as to Count 1 and, contingent on the grant of Ni substantive motion 1, as to Ni's proposed count 2 (Paper 29). Rauch opposes (Paper 53); Ni replies (Paper 59).

To the extent Ni substantive motion 2 is contingent upon the grant of Ni substantive motion 1, it is **dismissed** as moot because the contingency has not occurred.

As discussed above, the subject matter of Count 1 is directed to isolated DNA that encodes a polypeptide having an amino acid sequence that is at least 90% identical to SEQ ID NO:2 of Rauch's involved '642 patent, wherein the polypeptide binds TRAIL (FF 10 and 11). TRAIL is a member of the TNF ligand family and was known to be capable of inducing apoptosis (FF 27). To be accorded benefit of the filing date of an earlier filed application, the earlier application must provide a constructive reduction to practice of an embodiment within the count, i.e., a described and enabled anticipation of the subject matter of the count.

43. The '021 and '846 applications are both provisional applications.

44. The '021 application was filed 29 July 1997 (NX 2056, cover sheet).

45. The '846 application was filed 17 March 1997 (NX 2042, cover sheet).

46. Figure 1 of the '021 application is said to show the nucleotide and deduced amino acid sequences of "human Death Domain Containing Receptor 5" ("DR5") obtained from the cDNA clone deposited as ATCC Deposit No. 97920 on 7 March 1997 (NX 2056, p. 1, ll. 7-9; p. 6, ll. 5-6; p. 7, ll. 29-33; p. 9, ll. 9-12; p. 10, ll. 34-35).

47. According to the '021 specification, DR5 is a 411 amino acid protein (id., p. 26, ll. 9-10).

48. Example 6 of the '021 specification is said to show that a DR5 extracellular domain-Fc fusion construct ("DR5-Fc") binds TRAIL (id., p. 50, l. 6 - p. 51, l. 2; Figures 6A-6C).

49. Figure 1 of the '846 application is said to show the nucleotide and deduced amino acid sequences of DR5 obtained from the cDNA clone deposited as ATCC Deposit No. 97920 on 17 March 1997 (NX 2042, p. 1, ll. 5-6; 3, ll. 22-25; p. 5, ll. 24-27).

50. According to the '846 specification, DR5 is a 411 amino acid protein (id., p. 6, ll. 25-27).

51. Figure 2 of the '846 application is said to compare the deduced amino acid sequence of DR5 to the amino acid sequences of human tumor necrosis factor 1, human Fas protein and DR3 protein (id., p. 5, ll. 8-13).

52. According to the '846 specification, similarities between the amino acid sequences shown in Figure 2 "**strongly suggests** that DR5 is also a death domain containing receptor with the ability to induce apoptosis," i.e., that DR5 belongs to a subset of TNF-family receptors (id., p. 6, ll. 31-33, emphasis added).

53. Further according to the '846 specification, "TNF-family ligands induce various cellular responses by binding to TNF-family receptors, including the DR5 of the present invention. Cells which express the DR5 polypeptide **are believed to have** a potent cellular response to DR5 ligands ..." (NX 2042, p. 26, ll. 12-15, emphasis added).

54. The '846 specification defines a "TNF-family ligand" as a

> naturally occurring, recombinant, and synthetic ligands that are capable of binding to a member of the TNF receptor family and inducing the ligand/receptor signaling pathway. Members of the TNF ligand family include, but are not limited to, **DR5 ligands**, TRAIL, TNF-α, lymphdtoxin-α [sic] (LT-α, also known as TNF-ß), LT-ß (found in complex heterotrimer LT-α2-ß),

> FasL, CD40, CD27, CD30, 4-IBB, OX40 and nerve growth factor (NGF). [NX 2042, p. 31, ll. 4-9, emphasis added.]

55. The amino acid sequence of the DR5 protein shown in the respective Figures 1 of the '021 and '846 applications are identical.

56. It is undisputed that the amino acid sequence shown in Figures 1 of the '021 and '846 applications are at least about 93% identical to the amino acid sequence of SEQ ID NO:2 as recited in Count 1, with 411 of 440 total amino acids being identical (see Paper 53, p. 23 where Rauch admits Ni SMFs 7 and 8).

57. Thus, the '021 application describes an enabled embodiment within the scope of Count 1, i.e., a DNA sequence encoding a human Death Domain Containing Receptor 5 ("DR5") having an amino acid sequence that is at least 93% identical to the amino acid sequence of SEQ ID NO:2 of the '642 patent (FFs 46, 47 and 56) and which binds TRAIL (FF 48).

58. Rauch does not dispute Ni's claim to benefit for the purpose of priority of the filing date of its '021 application (Paper 53).

Based on the foregoing, we accord Ni benefit of the filing date of the '021 application as to Count 1.

While the '846 specification describes (Figure 1) an isolated DNA encoding polypeptide DR5 comprising a deduced amino acid sequence which is at least 90% identical to the amino acid sequence set forth in SEQ ID NO:2 of the '642 patent (411 of 440 amino acids are identical), the disclosure of the '846 application simply suggests that the DR5 polypeptide encoded by the isolated DNA is a death domain containing receptor with the ability to induce apoptosis

and suggests that the DR5 polypeptide binds a "DR5 ligand" (FFs 52 and 53). The disclosure of the '846 application does not describe preparing a DR5 polypeptide (or ligand binding portion thereof) or binding the ligand TRAIL to the DR5 polypeptide (or ligand binding portion thereof).

Ni's position is premised on classifying DR5 as a putative TNF "death receptor" protein based on the described similarity between DR5 and three previously known TNF death receptors TNFR1, Fas and DR3 in the '846 application. According to Ni, TNFR1, Fas and DR3 were all known to induce apoptosis upon activation and, therefore, that same function should be imputed to DR5 by virtue of the described similarities between the amino acid sequences of DR5 and the three death receptors. Ni argues that the '846 specification explicitly teaches that DR5 induces apoptosis and binds to a TNF ligand selected from a limited list including TRAIL. Ni further argues that, based on the doctrine of inherency, the '846 application need not expressly recite that DR5 binds TRAIL so long as the '846 application describes the subject matter of the count. [Paper 29, p. 2, ¶ 3 and ¶ bridging pp. 9-10.]

59. Ni relies on the direct testimony of John C. Reed, M.D., Ph.D. (NX 2099) in support of its position.

60. Dr. Reed has been qualified as an expert to give opinions on the subjects of apoptosis and of the tumor necrosis family of ligands (TNFs) and receptors (TNFRs), including death receptors.

61. According to Dr. Reed, the deduced amino acid sequence of human DR5 described in the '846 application has all the canonical (structural)

features of a classic death receptor of the TNFR family, i.e., a leader peptide, conserved cysteine-rich domain(s), a transmembrane domain and a cytosolic domain containing a "death domain" (NX 2099, ¶¶ 19-20 and 28-31).

62. Further according to Dr. Reed, the death domain "is necessary and sufficient for apoptosis induction, at least when overexpressed in mammalian cells" (NX 2099, ¶ 21).

63. Still further according to Dr. Reed, DR5 shares the highest degree of amino acid sequence identity with then known death receptor proteins human TNFR1, Fas and DR3 (NX 2099, ¶¶ 20 and 29).

64. Dr. Reed states that the deduced amino acid sequence of the "death domain" region of the DR5 protein described in Ni's '846 application was approximately 21, 32 and 33 percent identical to the amino acid sequences of the death domains of known death receptors Fas, TNFR1 and DR3, respectively, "using Lipman-Pearson Protein Alignment (with the following parameters: Ktuple 2; Gap Penalty 4; Gap Length Penalty 12)" (NX 2099, ¶ 31).

65. Dr. Reed opines that a death domain amino acid sequence identity of approximately 21-33 percent is "significant" because Chinnaniyan (NX 2058) reported that the death domain of DR3 was 47 and 23 percent identical to that of TNFR1 and Fas, respectively, while Marsters (NX 2059) reported that the death domain of DR3 was 48 and 20 percent identical to that of TNFR1 and Fas, respectively (NX 2099, ¶ 31).

66. Chinnaiyan reported using MegAlign™ software to align the compared amino acid sequences (NX 2058, Fig. 1).

67. MegAlign™ software can create alignments between two or more sequences according to different methods, e.g., the clustal method or the Jotun Hein method (see e.g., U.S. Patent 6,277,568, col. 8, ll. 22-41).

68. Neither Chinnayian nor Marsters reported the alignment program and parameters used to obtain their respective percent sequence identity scores.

69. Dr. Reed did not explain percent sequence identity scoring, e.g., how different alignment methods and parameters calculate percent sequence identity scores; how different alignment methods are compared (normalized to account for the use of parameter differences in sequence lengths, gaps, gap positions, etc.); the significance, if any, of comparing sequences within predicted structural features (e.g., a death domain or extracellular domain) versus over their entire primary amino acid sequence; standard error of the method(s) used; use of iteration, etc.

70. For example, according to Tartaglia,

> [i]t has been noted previously that the intracellular domain of TNF-R1 shares a weak homology (29% identity over 45 amino acids) with the intracellular domain of Fas antigen. Upon further inspection of these sequences, we noted that introduction of a 1 amino acid gap in the Fas sequence extended the region of homology an additional 20 amino acids (Figure 3). [NX 2067,[6] p. 846, col. 2, ¶ 1, emphasis added, citation omitted.]

[6] Tartaglia et al. (Tartaglia), "A Novel Domain within the 55 kd TNF Receptor Signals Cell Death," Cell, Vol. 74, pp. 845-853 (10 September 1993) (NX 2067).

71. Nonetheless, Dr. Reed believes that one of ordinary skill in the art would have reasonably expected the putative death receptor DR5 of the '846 specification to have utilities similar to known utilities of other known death receptors TNFR1, Fas and DR3 (NX 2099, ¶¶ 33-34).

72. According to Dr. Reed, "**the most reasonable conclusion to draw** from Ni's March 17, 1997 application is that DR5 is expected, by persons of ordinary skill in the art, to be a novel death receptor" and, therefore, skilled artisans "**would have predicted** that activation of DR5 would induce apoptosis" (NX 2099, ¶ 32, emphasis added).

73. According to Dr. Reed, induction of apoptosis involves activating (aggregating) the death receptor on the surface of a cell (in its membrane) and activating a family of caspase enzymes inside the cell (NX 2099, ¶ 24).

74. Further according to Dr. Reed, activation (aggregation) of the death receptor could be caused by (i) ligand binding to the death receptor, (ii) antibody binding to the death receptor or (iii) overexpression of the death receptor on the cell surface (NX 2099, ¶ 24).

75. Dr. Reed testified that

> if one would want to determine to which TNF ligand DR5 binds, Ni's March 17, 1997 application [i.e., the '846 application], in combination with what was known in the art at the time, provides all of the necessary information. For example, Ni's March 17, 1997 application states that **DR5 binds to a TNF-family ligand** (Exhibit 2042, pg. 4, ¶¶2-3; pg. 26, ¶1; pages 28-29; pg. 31, ¶1, pg. 31, ¶1 [sic]), which would have been expected by a person of ordinary skill in the art in view of the literature that was available by March 17, 1997. Additionally, Ni's March 17, 1997

> application specifically defines "a TNF family ligand" as a limited number of molecules, one of which is TRAIL. (Exhibit 2042, pg. 31, lines 4-9). The Ni March 17, 1997 application also teaches assays, such as cellular response **assays, that could be used to determine whether** TRAIL, or any other of **the listed TNF ligands, binds to DR5**. (Exhibit 2042, pg. 26, lines 12-26; pg. 27, line 21 through pg. 29, line 6). **Alternatively**, as of March 17, 1997, **it would have been routine** for one of ordinary skill in the art **to have tested whether DR5 binds to the TNF-family ligands recited in Ni's May [sic] 17, 1997 application**, including TRAIL. Thus, if one wanted to have determined whether DR5 bound to a TNF ligand, including TRAIL, the Ni March 17, 1997 application, in combination with what was known in the art at the time, teaches all of the needed information. [NX 2099, ¶ 56, emphasis and bracketed text added.]

76. Dr. Reed notes that while most TNF family receptors have been shown experimentally to bind to specific TNF family ligands, some receptors "do not have known ligands to date, or a delay of many years occurred before the specific ligand was established" (NX 2099, ¶ 18).

77. According to the '846 specification, there are eleven known TNF ligand members, i.e., TNF-α, lymphotoxin-α (LT-α, also known as TNF-ß), LT-ß (found in complex heterotrimer LT-α2-ß), FasL, CD40, CD27, CD30, 4-1BB, OX40 and nerve growth factor (NGF) (NX 2042, p. 1, ll. 21-25) and TRAIL (id., p. 31, ll. 6-9).

78. TRAIL was identified as a TNF family ligand by at least December 1995 (NX 2096).

79. The '846 specification defines "TNF-family ligand" as

> naturally occurring, recombinant, and synthetic ligands that are capable of binding to a member of the TNF receptor family and inducing the ligand/receptor

> signaling pathway. Members of the TNF ligand family include, but are not limited to, **DR5 ligands**, TRAIL, TNF-α, lymphdtoxin-α [sic] (LT-α, also known as TNF-ß), LT-ß (found in complex heterotrimer LT-α2-ß), FasL, CD40, CD27, CD30, 4-1BB, OX40 and nerve growth factor (NGF). [NX 2042, p. 31, ll. 4-9, emphasis added.]

80. Dr. Reed relies on Ni's later filed '201 application (NX 2056, Figure 6A) and on a later published August 1997 article (NX 2031[7]) to support his testimony that DR5 "necessarily" binds to TRAIL and "necessarily" induces apoptosis (NX 2099, ¶ 57).

A constructive reduction to practice requires a described and enabled anticipation under 35 U.S.C. § 102(g)(1). To fulfill the written description requirement, the patent specification must describe an invention in sufficient detail that one skilled in the art can clearly conclude that the inventor invented what is claimed. Lockwood v. Am. Airlines, Inc., 107 F.3d 1565, 1572, 41 USPQ2d 1961, 1966 (Fed. Cir. 1997). The specification "need not describe the claimed subject matter in exactly the same terms as used in the claims; it must simply indicate to persons skilled in the art that as of the [filing] date the applicant had invented what is now claimed." Eiselstein v. Frank, 52 F.3d 1035, 1038, 34 USPQ2d 1467, 1470 (Fed. Cir. 1995) (citing Vas-Cath Inc. v. Mahurkar, 935 F.2d 1555, 1562, 19 USPQ2d 1111, 1115 (Fed. Cir. 1991) and In re Wertheim, 541 F.2d 257, 265, 191 USPQ 90, 98 (CCPA 1976). Furthermore, "the fact that a characteristic is a necessary feature or result of a prior-art embodiment (**that is itself sufficiently described and enabled**) is enough for inherent anticipation,

[7] Guohua Pan, Jian **Ni**, Ying-Fei Wei, Guo-liang **Yu**, Reiner **Gentz**, Vishva M. Dixit, "An Antagonist Decoy Receptor and a Death Domain-Containing Receptor for TRAIL," Science, Vol. 277, pp. 815-18 (8 August 1997).

even if that fact was unknown at the time of the prior invention" (Toro Co. v. Deere & Co., 69 USPQ2d 1584, 1590 (Fed. Cir. 2004), bold emphasis added (citing Schering Corp. v. Geneva Pharmaceuticals, Inc., 339 F.3d 1373, 1378, 67 USPQ2d 1664, 1668-69 (Fed. Cir. 2003) and Atlas Powder Co. v. Ireco Inc., 190 F.3d 1342, 1347, 51 USPQ2d 1943, 1947-48 (Fed. Cir. 1999)).

Here, the subject matter of the count is directed to a family of DNA molecules which encodes a functional protein, i.e., a TNF death receptor protein that binds TRAIL (a TNF ligand known to be capable of inducing apoptosis). Relying on the testimony of Dr. Reed, Ni contends that the similarity of the deduced amino acid sequence of DR5 to the amino acid sequences of three known TNF death receptor proteins (TNFR1, Fas and DR3) as described in the '846 application is sufficient to characterize DR5 as a putative TNF death receptor protein and to reasonably predict that DR5 has utilities/functions similar to known death receptor proteins, e.g., induction of apoptosis upon activation.

Neither the disclosure of the '846 application nor the testimony of Dr. Reed is as explicit as Ni argues. The '846 application suggests that the protein encoded by the DNA of Figure 1 may be classified as a putative TNF death receptor protein. Dr. Reed testified that the most reasonable conclusion a person of ordinary skill in the art would draw from the '846 application is that DR5 "is expected ... to be a novel death receptor" (FF 64). However, the factual basis for this conclusion is not persuasive. The '846 specification does not describe preparing (e.g., expressing and purifying the product of the DNA of Figure 1) a DR5 polypeptide or ligand binding portion thereof. The '846 specification does

not describe an activated (functional) DR5 or identify the TNF ligand which activates (binds to) DR5.

Since TRAIL was known to be capable of inducing apoptosis (FF 27), identifying TRAIL as the TNF ligand which bound to DR5 in the '846 application would have been one way of describing DR5 as capable of inducing apoptosis. Dr. Reed testified that the '846 application "states that DR5 binds to a TNF-family ligand" and that there were "assays, that could be used to determine whether TRAIL, or any other of the listed TNF ligands, binds to DR5" (FF 75). Dr. Reed further testified that "it would have been routine for one of ordinary skill in the art to have tested whether DR5 binds to the TNF-family ligands recited" in the '846 application, "including TRAIL" (FF 75). Notably, the '846 specification enumerates "DR5 ligands" as separate and distinct ligands in the list of TNF ligands, including TRAIL (FF 54), the implication being that DR5 might bind to either a known TNF ligand, e.g., TRAIL, or an as yet unknown TNF ligand, i.e., a DR5 ligand, or another TNF ligand known to be capable of inducing another function, e.g., cell proliferation.

In short, there is neither explicit nor implicit disclosure in the '846 application said to show that the DR5 polypeptide encoded by the DNA of Figure 1 is a functional/bioactive protein. The cognate ligand for DR5 was not explicitly identified in the '846 application, although it would have been routine for one of ordinary skill in the art to do so using known techniques, as testified to by Dr. Reed (FF 75). However, Ni has failed to explain how such "routine experimentation" satisfies the written description requirement of 35 U.S.C. § 112,

first paragraph. Moreover, there could be no explicit description of an activated DR5 polypeptide based on antibody binding or overexpression in mammalian cells absent obtaining the DR5 polypeptide, e.g., by expressing (or overexpressing) the product of the DNA of Figure 1 to obtain a protein against which to raise an antibody. Furthermore, one of ordinary skill in the art could not have reasonably predicted the function(s) of DR5 based solely on the similarity between its deduced amino acid sequence as set forth in Figure 1 of the '846 application and the amino acid sequences of TNFR1, Fas and DR3 in view of the state of the art when the '846 application was filed for the following reasons.

Genes encode proteins by providing a sequence of nucleic acids that is translated into a sequence of amino acids. Methods used to identify novel genes are classified into two types, i.e., homology based or non-homology based. In homology based methods, for example, clones from a cDNA library are cloned and analyzed (sequenced). The resultant nucleotide sequences and/or deduced amino acid sequences are checked against databases for similarity (homology) to previously characterized sequences on the theory that molecules with similar sequences would be expected to perform similar functions. However, one of the difficulties in identifying a functional protein is that function depends not only on the amino acid sequence of the protein, but also on other factors, e.g., the structure of the protein.

In order for a protein to function properly its amino acid sequence (primary structure) must fold itself up into a complex three-dimensional shape which allows for molecular recognition. Molecular recognition often involves only a

small number of key amino acid residues on the functional surfaces of interacting molecules. These residues are dispersed in diverse regions of the primary amino acid sequence due to the complex structural organization of the protein. There are multiple levels to the structural organization of a protein. The *primary structure* of a protein refers to the linear arrangement of amino acid residues along a polypeptide chain. *Secondary structures* form through interactions between amino acids typically found near each other in the peptide chain which fold parts of the chain into regular structures, e.g., α helices and ß sheets. *Tertiary structure* folds both the secondary structures and the regions between them into compact three-dimensional shapes in an energetically favourable way. *Quaternary structure* refers to the organization of several polypeptide chains into a single protein molecule, e.g., hemoglobin is a tetramer. Consequently, amino acid residues rather near to each other in a protein's primary structure may be rather distant in the protein's ultimate quaternary structure. [See generally, MOLECULAR CELL BIOLOGY ("MCB"), second edition, Darnell et al., W.H. Freeman and Company, New York, NY (1990), pp. 44-48 (copy enclosed)].

For example, an enzyme is a protein that catalyzes a biochemical reaction. The function of an enzyme relies on the structure of its "active site," a specific cavity-like region on the surface of the three-dimensional enzyme which allows a snug fit (molecular recognition) between the enzyme and its substrate (reactant in the reaction being catalyzed). The active site contains key amino acids that bind the substrate and are involved in the reaction catalyzed by the enzyme. These key amino acids are brought into proximity (into the active site)

by protein folding. [See generally, MICROBIOLOGY: An Introduction, Tortora et al., The Benjamin/Cummings Publishing Company, Inc., Menlo Park, California, (1982) pp. 111-112, copy enclosed; MCB, pp. 55-65, copy enclosed.]

On the other hand, mutations that cause human disease often disrupt protein structure, thereby altering or abolishing normal protein function. For example, sickle cell anemia occurs in humans that are homozygous for a ß-hemoglobin gene that differs from the normal adult hemoglobin gene by a single base pair, resulting in a change in a single amino acid from glutamate to valine in position 5. This substitution is on the surface of the abnormal hemoglobin (Hb S) and changes the electrostatic charge on the surface of Hb S. When oxygen is removed from Hb S, the protein polymerizes into rigid crystals that deform a sickle cell patient's red blood cells. Thus, although normal hemoglobin and Hb S have virtually identical primary amino acid sequences, a single amino acid change in Hb S alters its quaternary structure and results in abnormal protein function. [See generally, CLINICAL DIAGNOSIS AND MANAGEMENT BY LABORATORY METHODS, sixteenth edition, J. B. Henry ed., W.B. Saunders Company, Philadephia (1979), Vol. I, p. 992, copy enclosed.]

Ergo, "[s]equence comparison can indicate whether an RNA or protein molecule or region of DNA is already known (identity) or has some degree of similarity to a known sequence" (MOLECULAR BIOLOGY AND BIOTECHNOLOGY, R. Meyers, ed., VCH Publishers, Inc., New York, NY (1995), p. 860, c. 1, ¶ 1, copy enclosed). However, since "[t]he function of nucleic acids

and proteins depend on their structure and involves complex interactions in three dimensions",

> [i]t is not presently understood whether it is possible, in general, to derive structure from sequence. Sequence alone is therefore often inadequate to determine function. Predictions made from sequence analysis need to be experimentally tested. Nonetheless, computer analysis of sequences is valuable in suggesting the most useful experiments to perform. [Id., p. 860, c. 1, ¶ 2.]

Indeed, the difficulties in predicting the structure and function of a protein from just its amino acid sequence (primary structure) are so well known in the art that the ability to characterize the structure and function of a protein from its amino acid sequence has been called the "Holy Grail" of molecular biology (RX 1061, p. 511, c.2, ¶ 1 to p. 512, c. 1, ¶ 1).

81. Genchong Cheng, Ph.D., is a witness for Rauch and has been qualified as an expert to give opinions on the subjects of signal transduction and gene expression networks through the TNFR, Toll-like receptor (TLR) and Nod receptor families during immune responses.

82. According to Dr. Cheng,

> [s]equence homology to other death domain-containing TNF receptors may be sufficient to convince one of ordinary skill in the art that a novel protein is a TNFR family member. However, sequence homology alone is not sufficient to support an assertion that a novel TNFR family member protein will induce specific biological activities such as apoptosis. Without additional data regarding the activity of a TNFR family member, such as, for example, the identity of the ligand with a known function (such as TRAIL) to which the receptor binds, one of ordinary skill in the art cannot reasonably predict the function of the TNFR family member. [RX 1049, ¶ 13.]

Ni's own witness, Dr. Reed, did not testify that the specification and figures of the '846 application would have reasonably conveyed to one of ordinary skill in the art that a DR5 having the deduced amino acid sequence shown in Figure 1 is in fact a functional death receptor protein based solely on its amino acid sequence (primary structure). Dr. Reed did not testify that one of ordinary skill in the art would have understood the '846 application to describe a DNA encoding a functional death receptor protein. Rather, Dr. Reed testified to "the most reasonable" (not the necessary and always) conclusion that one of ordinary skill in the art would have drawn from the disclosure of the '846 application (FF 72). Dr. Reed bases this conclusion on his testimony that there was "significant" percent sequence identity between the deduced amino acid sequence of DR5's death domain and the death domains of TNFR1, Fas and DR3 (FFs 64 and 65). However, we decline to credit that testimony because Dr. Reed did not provide a sufficient basis for his opinion. Dr. Reed did not explain how percent sequence identity scores were obtained, identify what alignment methods and parameters were used by the "references" (Chinnayian and Marsters (NX 2058 and NX 2059)), explain how percent identity scores based on different alignment methods and parameters relate to each other, what standard of error was typically found, whether iteration was necessary to obtain a statistically valid result, etc. 37 CFR § 41.158. Standing Order, ¶ 24. Further, as illustrated by the discussion of Hb S above, even very small differences between protein variants with highly similar amino acid sequences can produce significant differences in function.

Therefore, in view of the state of the art at the time the '846 application was filed and the testimony of both Drs. Reed and Cheng, we find that the '846 application does not describe an isolated DNA encoding a functional protein which binds TRAIL within the scope of Count 1. Rather, the '846 application describes an isolated DNA which encodes a polypeptide which may be preliminarily classified as a TNF death receptor protein based upon its deduced amino acid sequence. However, given the unpredictability of determining function from structure, a person skilled in the art would have had to carry out further research to identify the function(s) of the protein encoded by the DNA set forth in Figure 1 of the '846 application.

Anticipation is a question of fact, not a conclusion of law, no matter how reasonable that conclusion may appear to be. Putative assignment to a protein (sub)family does not assess the actual biological function/utility of a gene sequence and encoded protein product given the unpredictability of determining function from structure. Ni has failed to establish that the '846 application describes a DNA encoding a functional death receptor protein based solely on the disclosure of a deduced amino acid sequence. Brenner v. Manson, 383 U.S. 519, 532, 148 USPQ 689, 694 (1966) ("the presumption that adjacent homologues have the same utility has been challenged in the steroid field because of 'greater known unpredictability of compounds in that field.'").

Ni argues that the DR5 protein of the '846 application inherently binds TRAIL and that the '846 specification explicitly teaches that DR5 binds a TNF ligand selected from a limited list which includes TRAIL (Paper 29, p. 2, ¶ 3).

First, before considering whether a limitation is an inherent characteristic of an embodiment within the scope of a count, that embodiment must itself be sufficiently described and enabled. Toro, 69 USPQ2d at 1590. Thus, this argument fails because Ni has not established that the '846 application describes an enabled embodiment within the scope of Count 1 for the reasons above.

Second, arguing that DR5 binds a TNF ligand from a limited list which included TRAIL is also unpersuasive. The '846 specification does not explicitly identify TRAIL as the cognate ligand for DR5. The so-called "limited list" to which Ni refers apparently covers all the known and unknown ligands of the TNF family, i.e., the list enumerates the eleven then known TNF ligands and then adds a catch-all "DR5 ligands," seemingly in the event DR5 did not bind any of the then known TNF ligands. Neither the disclosure of the '846 application nor the testimony of Dr. Reed suggests that DR5 necessarily and always binds TRAIL or that DR5 binds a ligand selected from a limited subset of TNF ligands.

Third, while a specific DNA sequence may render a protein having a particular amino acid sequence obvious, a DNA is not a protein and, therefore, does not anticipate the encoded protein and its inherent properties.

Fourth, Ni's reliance on cited case law is misplaced. Ni argued that

> even without express appreciation of a limitation recited in a count, disclosure in a priority application of an embodiment which is later shown to *inherently* possess a characteristic satisfying that limitation is sufficient to establish constructive reduction to practice. *See e.g., Silvestri v. Grant*, 496 F.2d 593, 599, 181 U.S.P.Q. 706, 710 (CCPA 1974) ("The invention is not the language of the count but the subject matter thereby defined."); *See also Hudziak v. Ring*, 2005 Pat. App. LEXIS 26 (Bd. Pat. App. Intf., Sept. 2005) (confirming that a party's priority

> applications, which disclosed an antibody but did not state the antibody bound to a particular receptor protein (HER2) as recited in the count, were nonetheless constructive reductions to practice because subsequent evidence showed that the antibody bound HER2.) [Paper 29, p. 8, ¶ 1, original emphasis.]

Neither Silvestri nor Hudziak are on point. Silvestri has been discussed above (§ III. Ni Substantive Motion I). In Silvestri, the court held that the evidence established that Silvestri had prepared a new form of ampicillin, that Silvestri recognized and appreciated the existence of the new form of ampicillin and that the new form of ampicillin had utility. Id., 496 F.2d at 598-601, 181 USPQ at 709-712. The court acknowledged that the ampicillin of the count required a molecular weight of about 349 and greater storage stability than the previously known form of ampicillin. However, the court thought these were inherent properties of the new form of ampicillin that Silvestri was said to have obtained, recognized and described. Id., 496 F.2d at 599, 181 USPQ at 709. The court noted in Silvestri that the reduction to practice test does not require in haec verba appreciation of each of the limitations of the count:

> This standard does not require that Silvestri establish that he recognized the invention in the same terms as those recited in the count. The invention is not the language of the count but the subject matter thereby defined. Silvestri must establish that he recognized and appreciated as a new form, a compound corresponding to the compound defined by the count. Id., 496 F.2d at 599, 181 USPQ at 710.

Here, the DNA of the count is a precursor to a new compound, a protein which binds TRAIL. Thus, it is necessary to consider whether the '846 application describes the encoded protein and its properties/uses. While the '846

application describes a specific DNA, it only speculates that the DNA of Figure 1 encodes a protein having the desired properties. Ni is not in the same position as Silvestri, whose application specifically described an ampicillin compound, specifically recognized it as a new form of ampicillin and specifically described certain properties of the compound. Ni's application describes a precursor to an encoded protein, but only speculates on the nature and properties of that protein. Therefore, Silvestri is not on point.

Similarly, in Hudziak v. Ring, 80 USPQ2d 1018, 1019 (Bd. Pat. App. & Int. 2005), the count was directed to a monoclonal antibody that bound human epidermal growth factor receptor 2 (HER2). A panel of the Board decided that Chiron's (Ring's real party-in-interest) 1984 application disclosed an embodiment within the count, i.e., a murine monoclonal antibody designated 454C11. Id. The panel noted that the 1984 application (06/577,976) stated that hybridomas which produced 454C11 were deposited with the ATCC and that evidence submitted by Chiron established that 454C11 bound HER2. Id. at 1020-21.

83. The panel also noted in its decision (Paper 258, p. 129) that "Table 3 of the 1984 application reports the binding of antibodies to breast cancer cell lines and indicates that 454C11 binds to SKBR3 cells, which are now known to express HER2. (CX 1081, p. 3)."

Thus, in Hudziak, Chiron was said to have actually prepared an embodiment within the count, monoclonal antibody 454C11, and to have described it as a new protein and appreciated one of its properties/functions, i.e., that it bound to breast cancer cells. Ni's '846 application, on the other hand, describes a

precursor to an encoded protein, but only speculates on the nature and properties of that protein. Therefore, Hudziak is not on point.

Since Ni has failed to establish that '846 application describes an enabled death receptor protein of the TNF family, we do not reach the issue of what the inherent characteristics of that protein are. In both Silvestri and Hudziak, the application specifically described compounds that were recognized as novel and as having certain properties. These described and characterized compounds were later found to have other properties required by the count. Here, Ni's '846 application does not describe and characterize the protein encoded by the DNA of Figure 1. Ni's application only speculates on the nature and properties of the encoded protein and that speculation is insufficient to show possession of an enabled embodiment within the count which is later found to have other properties required by the count.

84. Lastly, Ni argues DR5 DNA might be used as "diagnostic reagents for detecting mutated forms of DR5 associated with a dysfunction (*e.g.*, diseases which result from under-expression, over-expression or altered expression of DR5, such as tumors or autoimmune diseases)" or "in gene transfer applications" (Paper 29, ¶ bridging pp. 10-11).

As noted by Rauch in its opposition (Paper 53, pp. 13-14), these uses are premised on expression of the encoded DR5 protein being linked to a particular disease state or on the ability of DR5 to induce apoptosis. The '846 application does not describe how to use a DNA encoding DR5 because it only speculates on the nature and properties of DR5.

Based on the foregoing, Ni is not entitled to benefit of the filing date of the '846 application as to Count 1.

In conclusion, Ni substantive motion 2 is **granted-in-part**, **denied-in-part** and **dismissed-in-part**.

VII. Rauch Substantive Motion 2

Pursuant to 37 CFR § 41.121(a)(1)(i), Rauch moves to redefine the scope of the interference by designating Ni claims 321, 322, 324 and 477 of the '583 application as corresponding to Count 1 (Paper 34). Ni opposes (Paper 49); Rauch replies (Paper 61).

"A claim corresponds to a count if the subject matter of the count, treated as prior art to the claim, would have anticipated or rendered obvious the subject matter of the claim." 37 CFR § 1.207(b)(2). The subject matter of Count 1 is directed to a genus of isolated DNAs that encode a polypeptide having an amino acid sequence that is at least 90% identical to SEQ ID NO:2 of Rauch's involved '642 patent, wherein the polypeptide binds TRAIL (FF 10 and 11). It is undisputed that the TRAIL-R protein described in SEQ ID NO:2 of the '642 patent is the 440 amino acid isoform of a receptor protein alternatively referred to in the literature as TR-2, DR5, Apo-2 TRICK2 and KILLER (FF 23).

85. It is undisputed that the DR5 protein disclosed in the '583 application is the 411 amino acid isoform of TR-2 (see Paper 49, p. B-1, where Ni admits Rauch SMFs 6 and 7 as set forth in Paper 34, p. 10).

86. Amino acid residues 1 to 440 of SEQ ID NO:2 of Rauch's '642 patent are identical to amino acid residues -51 to 360 of SEQ ID NO:2 of Ni's '583

application except for the inclusion of additional amino acid residues 185 to 213 in SEQ ID NO:2 of Rauch's '642 patent (RX 1044, pp. 58-60; RX 1046, ccs. 33-35).

87. Amino acid residues 52-184 of SEQ ID NO:2 of Ni's 583 application constitute the extracellular domain of DR5, which is the region of DR5 said to bind TRAIL (RX 1004, p. 6, ll. 6-7 and Example 6, pp. 53-54).

88. Claim 319 of the '583 application reads:

> An isolated polynucleotide comprising a nucleic acid which encodes amino acids 1 to 360 of SEQ ID NO:2.

89. Claim 321 of the '583 application reads:

> The polynucleotide of claim 319, wherein said nucleic acid encodes amino acids -50 to 360 of SEQ ID NO:2.

90. Amino acid residues 2 to 440 of SEQ ID NO:2 of Rauch's '642 patent are identical to amino acid residues -50 to 360 of SEQ ID NO:2 of Ni's '583 application except for the inclusion of additional amino acid residues 185 to 213 in SEQ ID NO:2 of Rauch's '642 patent (RX 1044, pp. 58-60; RX 1046, ccs. 33-35).

91. The polynucleotide of claim 321 includes the region of DNA which encodes the extracellular domain of DR5 and, therefore, encodes a polypeptide that binds TRAIL.

92. Therefore, the polynucleotide of Ni claim 321 encodes a polypeptide which is at least 90% identical to the amino acid sequence of SEQ ID NO:2 of Rauch's '642 patent and which binds TRAIL.

93. Claim 322 of the '583 application reads:

> The polynucleotide of claim 321, which comprises nucleotides 133 to 1362 of SEQ ID NO:1.

94. Nucleotide residues 133 to 1362 of SEQ ID NO:1 of the '583 application encode amino acid residues -50 to 360 of SEQ ID NO:2 of the '583 application which are identical to amino acid residues 2 to 440 of SEQ ID NO:2 of Rauch's '642 patent except for the inclusion of additional amino acid residues 185 to 213 in SEQ ID NO:2 of Rauch's '642 patent (RX 1044, pp. 58-60; RX 1046, ccs. 33-35).

95. The polynucleotide of claim 322 includes the region of DNA which encodes the extracellular domain of DR5 and, therefore, encodes a polypeptide that binds TRAIL.

96. Therefore, the polynucleotide of Ni claim 322 encodes a polypeptide which is at least 90% identical to the amino acid sequence of SEQ ID NO:2 of Rauch's '642 patent and which binds TRAIL.

97. Claim 324 of the '583 application reads:

> The polynucleotide of claim 322, which comprises nucleotides 130 to 1362 of SEQ ID NO:1.

98. Nucleotide residues 130 to 1362 of SEQ ID NO:1 of the '583 application encode amino acid residues -51 to 360 of SEQ ID NO:2 of the '583 application which are identical to amino acid residues 1 to 440 of SEQ ID NO:2 of Rauch's '642 patent except for the inclusion of additional amino acid residues 185 to 213 in SEQ ID NO:2 of Rauch's '642 patent (RX 1044, pp. 58-60; RX 1046, ccs. 33-35).

99. The polynucleotide of claim 324 includes the region of DNA which encodes the extracellular domain of DR5 and, therefore, encodes a polypeptide that binds TRAIL.

100. Therefore, the polynucleotide of Ni claim 324 encodes a polypeptide which is at least 90% identical to the amino acid sequence of SEQ ID NO:2 of Rauch's '642 patent and which binds TRAIL.

101. Claim 476 of the '583 application reads:

> An isolated polynucleotide comprising a nucleic acid which encodes the mature amino acid sequence encoded by the cDNA clone in ATCC Deposit No. 97920.

102. Claim 477 of the '583 application reads:

> The polynucleotide of claim 476, wherein said nucleic acid encodes the complete amino acid sequence encoded by the cDNA clone in ATCC Deposit No. 97920.

103. According to the '583 specification, the cDNA in ATCC Deposit No. 97920 encodes a DR5 polypeptide having the amino acid sequence set forth in SEQ ID NO:2 (RX 1044, p. 9, ll. 5-8).

104. The polynucleotide of claim 477 includes the region of DNA which encodes the extracellular domain of DR5 and, therefore, encodes a polypeptide which binds TRAIL.

105. Therefore, the polynucleotide of Ni claim 477 encodes a polypeptide which is at least 90% identical to the amino acid sequence of SEQ ID NO:2 of Rauch's '642 patent and which binds TRAIL.

106. In essence, Rauch's position is that "as long as a single species of a claim falls within the count, then that claim corresponds to the count" (Paper 34, p. 5, ¶ 1).

Rauch has established that each species of isolated polynucleotide recited in Ni claims 321, 322, 324 and 477 falls within the generic isolated TRAIL-R DNA

of Count 1 (FF 73-92). Rauch's position is that "if Count 1 were prior art to the Ni claims, it would anticipate the claims" (Paper 34, p. 5, ¶ 2).

A prior art species within a claimed genus reads on the generic claim and anticipates. In re Gostelli, 872 F.2d 1008, 1010, 10 USPQ2d 1614, 1616 (Fed. Cir. 1989). However, a species claim is not necessarily obvious in light of a prior art disclosure of a genus. In re Baird, 16 F.3d 380, 382, 29 USPQ2d 1550, 1552 (Fed. Cir. 1994). In other words, the "earlier disclosure of a genus does not necessarily prevent patenting a species member of that genus." Eli Lilly & Co. v. Bd. of Regents of the Univ. of Washington, 334 F.3d 1264, 1270, 67 USPQ2d 1161, 1165 (Fed. Cir. 2003)(citing Bristol-Myers Squibb Co. v. Ben Venue Labs., Inc., 246 F.3d 1368, 1380, 58 USPQ2d 1508, 1516-17(Fed. Cir. 2001)).

Here, Rauch has the burden of establishing that each of Ni claims 321, 322, 324 and 477 would have been anticipated or rendered obvious by the subject matter of Count 1. 37 CFR § 41.121(b). Simply showing that a species claim falls within the subject matter of a generic count (see Paper 34, pp. 5-7) does not suffice to establish that the claim is anticipated or rendered obvious by the subject matter of the count. Rauch has not established why any of Ni claims 321, 322, 324 and 477 would be unpatentable over the subject matter of Count 1, i.e., why each of these claims is an obvious species within the generic subject matter of the count. Therefore, Rauch has failed to meet its burden.

Based on the foregoing, Rauch substantive motion 2 is **denied**.

VIII. Rauch Substantive Motion 3

Pursuant to 37 CFR § 41.121(a)(1)(iii) and the Order issued 29 November 2005 (Paper 26), Rauch moves for judgment that Ni claims 287, 289-299, 351-361, 389-403, 431, 432, 434-442, 446, 448-458, 507-517, 581 and 623-632 ("Ni claims at issue") are unpatentable under 35 U.S.C. §§ 102(a) and/or (e) as clearly anticipated by one or more of U.S. Patent 6,642,358 ("the '358 patent," RX 1042), U.S. Patent 6,072,047 ("the '047 patent," RX 1048), U.S. Patent 6,569,642 ("the '642 patent," RX 1046) and WO 98/35986 ("WO '986," RX 1032) (collectively, "the Rauch references") (Paper 35, ¶ bridging pp. 2-3). Ni opposes (Paper 50); Rauch replies (Paper 63).

107. The '358 patent issued 4 November 2003, based on application 09/578,392, filed 25 May 2000, which is a divisional of application 08/883,036, filed 26 June 1997, which is a continuation-in-part of application 08/869,852, filed 4 June 1997, which is a continuation-in-part of application 08/829,536, filed 28 March 1997, which is a continuation-in-part of application 08/815,255, filed 12 March 1997, which is a continuation-in-part of application 08/799,861, filed 13 February 1997 (RX 1042, title page).

108. The '047 patent issued 6 June 2000 based on application 08/883,036, filed 26 June 1997, which is a continuation-in-part of application 08/869,852, filed 4 June 1997, which is a continuation-in-part of application 08/829,536, filed 28 March 1997, which is a continuation-in-part of application 08/815,255, filed 12 March 1997, which is a

continuation-in-part of application 08/799,861, filed 13 February 1997 (RX 1048, title page).

109. The '642 patent issued 27 May 2003, based on application 09/536,201, filed 27 March 2000, which is a continuation of application 08/883,036, filed 26 June 1997, which is a continuation-in-part of application 08/869,852, filed 4 June 1997, which is a continuation-in-part of application 08/829,536, filed 28 March 1997, which is a continuation-in-part of application 08/815,255, filed 12 March 1997, which is a continuation-in-part of application 08/799,861, filed 13 February 1997 (RX 1046, title page).

110. WO '986 published 20 August 1998, based on international application PCT/US98/02239, filed 11 February 1998 (RX 1032, title page).

According to the relevant paragraphs of 35 U.S.C. § 102:

> A person shall be entitled to a patent unless--
>
> (a) the invention was known or used by others in this country, or patented or described in a printed publication in this or a foreign country before the invention thereof by the applicant for patent, or
>
> * * * * *
>
> (e) the invention was described in (1) an application for a patent, published under section 122(b), by another filed in the United States before the invention by the applicant for patent or (2) a patent granted on an application for patent by another filed in the United States before the invention by the applicant for patent, except that an international application filed under the treaty defined in section 351(a) shall have the effects for the purposes of this subsection of an application filed in the United States only if the international

> application designated the United States and was published under Article 21(2) of such treaty in the English language, or
>
> * * * * *

References based on international applications that were filed prior to 29 November 2000 are subject to the former version of 35 U.S.C. § 102(e),[8] i.e.,

> A person shall be entitled to a patent unless --
>
> (e) the invention was described in a patent granted on an application for patent by another filed in the United States before the invention thereof by the applicant for patent, or on an international application by another who has fulfilled the requirements of paragraphs (1), (2), and (4) of section 371(c) of this title before the invention thereof by the applicant for patent.

A prima facie case is made out under § 102(a) if, within a year of the filing date, the invention, or an obvious variant thereof, is described in a "printed publication" whose authorship differs from the inventive entity unless it is stated within the publication itself that the publication is describing the applicant's work. In re Katz, 687 F.2d 450, 215 USPQ 14 (CCPA 1982).

111. None of the Rauch references issued or published prior to the 17 March 1998 filing date of Ni's claims at issue.[9]

112. None of the Rauch references qualify as prior art under § 102(a) vis-à-vis Ni's claims at issue.

Therefore, to the extent Rauch substantive motion 3 seeks a judgment that any of Ni claims at issue are unpatentable under § 102(a) over any of the Rauch

[8] Pursuant to § 13205 of Pub. L. 107-273.
[9] Rauch has not argued prior public knowledge or use of the subject matter of any of Ni's claims at issue.

references, the motion is **denied**. We now consider whether any of the Rauch references qualify as prior art under § 102(e).

WO '986 is based on an international application filed prior to 29 November 2000 (FF 110). Therefore, it must satisfy the requirements of then applicable former § 102(e) in order to qualify as prior art. Rauch has neither argued nor shown that WO '986 satisfies the requirements of the applicable § 102(e) (see Paper 35, p. 19, ¶ 1). Thus, Rauch has not established that WO '986 qualifies as prior art under the applicable § 102(e) vis-à-vis Ni's claims at issue. Consequently, to the extent Rauch substantive motion 3 seeks a judgment that any of Ni claims at issue are unpatentable under § 102(e) as anticipated by WO '986, the motion is **denied**.

As indicated above (FF 107-109), the '358, '047 and '642 patents are related. The '047 patent issued based on application 08/833,036 and the '358 and '642 patents issued based on an application identified as a divisional or a continuation, respectively, of application 08/883,036, filed on 26 June 1997. The filing date of the 08/833,036 application is prior to the 17 March 1998 filing date of Ni's involved claims and prima facie qualifies as prior art under § 102(e) against Ni's claims at issue. It is not necessary to consider whether Ni's claims at issue are anticipated by the '358 and '642 patents, if Ni's claims are anticipated by the '047 patent.

Claim chart appendix G attached to Rauch substantive motion 3 correlates the disclosure of the '047 patent to each of the limitations of each of Ni's claims at issue. Therefore, Rauch substantive motion 3, when considered in light of the

evidence relied upon in support of the motion, establishes a sufficient basis for holding Ni's claims at issue prima facie unpatentable under 35 U.S.C. § 102(e) as anticipated by the '047 patent.

113. Ni does not contest that the '047 patent describes the subject matter of its claims at issue.

114. Rather, Ni contends that the '047 patent does not qualify as prior art because Ni's '583 application claims are argued to be entitled to benefit of the 17 March 1997 filing date of its '846 provisional application (Paper 50, pp. 3-4; p. 10. ¶ 1; and Appendix E).[10]

115. Rauch maintains that Ni cannot obtain benefit of its '846 application due to a lack of utility (Paper 35, p. 2, ¶ 3 and p. 19, ¶ 2 - p. 21, ¶ 1).

As stated in In re Fisher, 421 F.3d 1365, 1378 USPQ2d 1225, 1235 (Fed. Cir. 2005),

> [i]t is well established that the enablement requirement of § 112 incorporates the utility requirement of § 101. The how to use prong of section 112 incorporates as a matter of law the requirement of 35 U.S.C. § 101 that the specification disclose as a matter of fact a practical utility for the invention. If the application fails as a matter of fact to satisfy 35 U.S.C. § 101, then the application also fails as a matter of law to enable one of ordinary skill in the art to use the invention under 35 U.S.C. § 112.

The salient question is whether Ni's claims at issue are entitled to benefit of the 17 March 1997 filing date of Ni's '846 provisional application, thereby antedating the 26 June 1997 filing date of the '047 patent. Benefit for purposes

[10] We need not consider whether Ni's '583 application claims are entitled to § 119(e) benefit of the 29 July 1997 filing date of Ni's '021 provisional application because the 29 July 1997 is after the 26 June 1997 filing date of the 08/833,036 application which issued as Rauch's '047 patent.

of antedating prior art, in this case, benefit under 35 U.S.C. § 119(e), is different from benefit for the purpose of priority. As set forth in section 119(e) of Title 35 of the United States Code,

> [a]n application for a patent filed under section 111(a) or section 363 of this title for an invention disclosed in the manner provided by the first paragraph of section 112 of this title in a provisional application filed under section 111(b) of this title, by an inventor or inventors named in the provisional application, shall have the same effect, as to such invention, as though filed on the date of the provisional application filed under section 111(b) of this title, if the application for patent filed under section 111(a) or section 363 of this title is filed not later than 12 months after the date on which the provisional application was filed and if contains or is amended to contain a specific reference to the provisional application.

116. Ni's '583 application claims benefit under § 119(e) of its '846 provisional application filed 17 March 1997 (RX 1044, p. 83).

The parties disagree whether the disclosure of Ni's '846 application satisfies the description and enablement requirements of § 112, first paragraph, as to the full scope of the subject matter of Ni's '583 application claims at issue.

Ni cites to specific disclosures in its '846 application said to describe every element of its claims at issue (Appendix E attached to Ni opposition 3 to Rauch motion 3). Ni argues that the '846 application discloses that DR5 DNA can be used to express (produce) DR5 polypeptides, which in turn, can be used (a) to make anti-DR5 antibodies for treating diseases associated with apoptosis or (b) as diagnostic reagents for detecting mutated forms of DR5 associated with a dysfunction or (c) as antagonists of DR signaling (Paper 50, pp. 10-13). Ni further argues that the DR5 DNA molecule itself can be used (d) as diagnostic

reagents to detect mutated forms of DR5 associated with a dysfunction, e.g., diseases which result from under-expression, over-expression or altered expression of DR5 or (e) in gene transfer applications to increase cellular apoptosis (Paper 50, pp. 13-14). Ni still further argues that DR5 antisense DNA or RNA can be used to inhibit DR5 gene expression (Paper 50, p. 14).

117. Dr. Reed, testifying for Ni, stated that the technology necessary to achieve these functions was within routine skill in the art, e.g., a skilled artisan would know how to express and purify a protein (e.g., DR5) from cDNA (e.g., DR5 cDNA), how to produce antibodies that bind a desired protein (e.g., DR5), etc. (e.g., NX 2099, ¶ 35-46).

118. Dr. Reed further testified that the uses for DR5 DNA and its encoded DR5 polypeptide described in the '846 application would have been believable to one of ordinary skill in the art because the asserted uses had previously been shown to be recognized uses of TNF death receptors TNFR1, Fas and/or DR3 (NX 2099, 33-34, 47-52).

Essentially, Dr. Reed's testimony as to the utility/enablement of DR5 DNA or its encoded DNA polypeptide is based on the assumption that the DR5 DNA described in the '846 application encoded a functional TNF death receptor and, therefore, what was known about the use of other death receptors was applicable to DR5 DNA and its encoded protein (see e.g., NX 2099, ¶¶ 49 and 50 ("[b]ased on precedent from prior work in the field of TNF-family receptors" and "[b]ased on precedent from the literature where agonistic and antagonistic antibodies to other TNF-family death receptors had been produced and characterized,"

respectively)). According to Ni, Dr. Reed "has testified unequivocally that 'you can reasonably make a prediction based on homology alone" and by analyzing "the particular subfamily of proteins to which DR5 belongs, *i.e.,* death receptors", "the most reasonable conclusion to draw from Ni's March 17, 1997 application is that DR5 is expected, by persons of ordinary skill in the art to be a novel death receptor [and that] a person of ordinary skill in the art would have predicted that activation of DR5 would induce apoptosis" (Paper 50, p. 16, ¶ 2, citations omitted). The disclosure cited by Ni in its Appendix E is no more specific than Dr. Reed's testimony. For example, in the second paragraph of the third column on page 2 of Appendix E, Ni points to page 6, lines 28-32 of the '846 application as disclosing that "[t]he homology DR5 shows to other death domain containing receptors strongly indicates that DR5 is also a death domain containing receptor with the ability to induce apoptosis." According to Ni, Dr. Reed properly focused on the subset of known death receptors and the "single" function that unites them, i.e., their ability to induce apoptosis (Paper 50, pp. 15-16).

Rauch, on the other hand, argues that sequence homology alone is insufficient to establish that the encoded DR5 polypeptide disclosed in the '846 application is in fact a TNF family death domain receptor. According to Rauch, unless the disclosure of the '846 application shows DR5 to be an actual TNF family member receptor, e.g., by identification of a known TNF ligand as its cognate ligand or by specific experimental data showing that DR5 induces a TNFR-mediated biological activity, e.g., apoptosis, inflammatory response, etc.,

the '846 application fails to disclose a specific, substantial and credible utility for the encoded protein and, therefore, for the claims at issue.

119. Dr. Cheng, testifying for Rauch, stated that Ni's '846 application

> discloses the DNA and amino acid sequence of the 411 amino acid isoform of TR-2, which they refer to as DR5. DR5 was identified based on sequence homology to other death domain-containing members of the TNFR family, including TNFR-1, DR3, and Fas ('846 Provisional, page 5, lines 21-24). The applicants assert that agonists to DR5 can be used to increase apoptosis, while antagonists to DR5 can be used to inhibit apoptosis. This assertion is based entirely on sequence homology between DR5 and the death domain-containing receptors TNFR-1, DR3, and Fas. However, the '846 Provisional does not identify a ligand for DR5 and does not contain any experimental data regarding DR5 function.
>
> Sequence homology to other death domain-containing receptors may be sufficient to convince one of ordinary skill in the art that a novel protein is a TNFR family member. However, sequence homology alone is not sufficient to support an assertion that a novel TNFR family member will induce specific biological activities such as apoptosis. Without additional data regarding the activity of a TNFR family member, such as, for example, the identity of the ligand with a known function (such as TRAIL) to which the receptor binds, one of ordinary skill in the art cannot reasonably predict the function of the TNFR family member. This is because TNFR family members are involved in complex signal transduction pathways, which can affect a wide spectrum of biological activities including apoptosis, inflammatory response, cell proliferation, cell survival, and other activities. The binding of certain TNFR family members by their corresponding ligands can lead to activation of multiple signal transduction pathways. As stated above, Ni's '846 Provisional contains no data regarding the ligand for DR5, nor does it disclose experimental data of its function. Without knowing more information about the activity of DR5, such as for example its specificity for a ligand with a known function, one of ordinary skill in the art could not

> reasonably predict the function of the TNFR family member. [RX 1049, ¶¶ 12-13.]

For essentially the reasons set forth in our analysis in "§ VI. Ni Substantive Motion 2" above, we credit the testimony of Dr. Cheng over that of Dr. Reed. In short, one of ordinary skill in the art might classify the product encoded by the DNA set forth in Figure 1 of the '846 application as a possible TNF death receptor protein based on the deduced amino acid sequence of the product. However, given the unpredictability of determining function from structure (the "Holy Grail" of molecular biology), a skilled artisan would have had to carry out further research to identify the function(s) of the protein encoded by the DNA set forth in Figure 1 of the '846 application. Thus, the disclosure of the '846 application fails to satisfy the "how-to-use" requirement of § 112, first paragraph, as to the subject matter of the Ni's claims at issue. Ni's claims at issue are not entitled to § 119(e) benefit of the filing date of the '846 application and the '047 patent still qualifies as prior art under § 102(e). Therefore, Ni claims 287, 289-299, 351-361, 389-403, 431, 432, 434-442, 446, 448-458, 507-517, 581 and 623-632 ("Ni's claims at issue") are unpatentable under 35 U.S.C. § 102(e) as clearly anticipated by U.S. Patent 6,072,047. It is not necessary to consider whether Ni's claims at issue are also anticipated by either patent '358 or '642.

In its opposition, Ni argues that Rauch substantive motion 3 should be denied on procedural grounds because it does not seek judgment that all of Ni's corresponding claims are unpatentable and, therefore, is not a proper threshold motion (Paper 50, p. 2, ¶ 3 - p. 2, ¶ 2; p. 7, ¶ 2 - p. 9, ¶ 3). Rauch substantive motion 3 is an ordinary attack on patentability. Ni has not provided any basis

requiring a motion for unpatentability to attack all of a party's involved claims and we know of none. Therefore, this argument is without merit.

Based on the foregoing, Rauch substantive motion 3 is **granted to the extent** that Ni's claims 287, 289-299, 351-361, 389-403, 431, 432, 434-442, 446, 448-458, 507-517, 581 and 623-632 are unpatentable under 35 U.S.C. § 102(e) as clearly anticipated by U.S. Patent 6,072,047 and **otherwise denied**.

IX. Rauch Miscellaneous Motion 5

Pursuant to 37 CFR § 41.115(c), Rauch seeks to exclude selected portions of the direct testimony of Dr. Reed that reference a person of ordinary skill in the art from evidence (NX 2099, ¶¶ 16, 19, 21-28, 30-43, 45-48, 50-52, 56, 63 and 64), contending that his definition of ordinary skill "is so broad that it fails to limit 'one of ordinary skill in the art' to any substantive or realistic meaning of such person" (Paper 71, ¶ bridging pp. 4-5). Rauch further seeks to exclude selected portions of the redirect testimony of Dr. Reed from evidence as leading and prejudicial (NX 2123, p. 172, l. 20; p. 173, ll. 7-8; p. 172, l. 25 through p. 173, l. 2) (id., p. 8, ¶¶ 2-3). Ni opposes (Paper 75); Rauch replies (Paper 83).

120. Rauch timely filed its objections to evidence sought to be excluded (RX 1093 and NX 2123, p. 172, l. 20 and p. 173, ll. 7-8).

Rauch identifies the objected to testimony of Dr. Reed as submitted in support of Ni substantive motion 2, Ni reply 2 and Ni opposition 3 to Rauch substantive motion 3 (Paper 71, Appendix D). First, Rauch's arguments go to the weight to be accorded Dr. Reed's testimony, not to its admissibility. Second, having considered the testimony of both Dr. Reed and Dr. Cheng, we credited

the testimony of the latter over that of the former as discussed in our denial of the relevant portion of Ni substantive motion 2 and in our granting of the relevant portion of Rauch substantive motion 3. Therefore, Rauch substantive motion 5 is **dismissed** as moot since we have not relied upon either the direct or redirect testimony of Dr. Reed to Rauch's detriment.

X. Ni Miscellaneous Motion 4

Pursuant to 37 CFR § 41.155(c), Ni seeks to exclude from evidence:

(a) exhibits related to Rauch's priority statements in (i) related interference 105,240 (RX 1025 and RX 1038), (ii) this interference (RX 1051) and (iii) related interference 105,381 (RX 1052 and RX 1054);

(b) direct (RX 1074) and deposition testimony (NX 2179-2181) of Dr. Gavin R. Scranton in related interference 105,240;

(c) direct testimony of Norman Boiani (RX 1075); and

(d) selected portions of the deposition testimony of Dr. Cheng (NX 2124, p. 132, l. 16 through p. 135, l. 5; p. 135, l. 9 through p. 136, l. 13 (with errata sheets)) (Paper 80). Rauch opposes (Paper 76); Ni replies (Paper 82).

Ni contends (Paper 80, pp. 22-23) that

> RX 1025, RX 1038, RS 1051, RX 1052 and RX 1054 should be excluded under FRE 901 for lack of authentication and lack of foundation. In addition, these exhibits should be excluded under FRE 1001 (4), 1002, and 1003, *inter alia*, because none of these exhibits appear to be originals nor admissible duplicates of the originals. Furthermore, these exhibits should be excluded under FRE 403, *inter alia*, because its probative value, if any, is outweighed by considerations of waste of time, lack of authentication and the reliability of the copies.

> Furthermore, RX 1074, the declaration of Dr. Gavin R. Screaton, should be excluded under FRE 403 because its probative value, if any, is outweighed by confusion of the issues. In addition, RX 1074 should be excluded under 37 C.F.R. § 41.122(b) because the declaration does not respond to arguments raised in an opposition but merely is an attempt by Rauch to make additional arguments in a reply that should have been raised in a motions. Furthermore, contingent upon the Board excluding RX 1074, Party Ni moves to exclude NX 2179, NX 2080 and NX 2181 for being irrelevant under FRE 401 and confusing the issues under FRE 403.
>
> In addition, Party Ni moves to exclude RX 1075, the Declaration of Norman Boiani under FRE 1002 because Exhibit A appears to be a photocopy, not an original, of a laboratory notebook page. Furthermore, Party Ni moves to exclude RX 1075 under FRE 403 because Exhibit A is taken out of context of the rest of the laboratory notebook. Party Ni's inability to determine the context of Exhibit A is unfairly prejudicial and this prejudice far outweighs any probative value of RX 1075.
>
> Lastly, the above-cited portions of NX 2124 should be excluded under FRE 611(c), FRE 403, and Cross Examination Guideline [3] of the Standing Order. The leading questions asked by Rauch's counsel clearly suggested single answers to the witness which resulted in the interjection of the opinions of counsel for Rauch in place of Dr. Cheng's opinions. Thus, the prejudicial effect of the cited testimony far outweighs its probative value, and the above-cited evidence should be excluded or, at most, accorded little weight by the Board.

Ni's motion has serious procedural defects. Rule 155(c) provides that a motion to exclude evidence must explain the objections and identify the objections in the record in order. As explained in Standing Order ¶ 21.3(a) a motion to exclude evidence shall (1) identify where in the record the objection was originally made and (2) identify where in the record the evidence was relied

upon by the opponent, and (3) address objections to exhibits (in whole or in part) in exhibit numerical order. According to Standing Order ¶ 21.1, the objection to the admissibility of evidence should be filed as part of a motion to exclude the evidence.

121. Ni contends that it timely objected to exhibits RX 1025, RX 1038, RX 1051, RX 1052 and RX 1054 as shown in exhibits NX 2192 and NX 2193, filed in support of its motion.

122. Ni exhibits NX 2192 and NX 2193 are "REDACTED" papers entitled "NI OBJECTIONS TO THE ADMISSIBILITY OF RAUCH'S SUPPLEMENTAL EXHIBIT 1054 AND RAUCH'S RESPONSES TO NI'S OBJECTIONS TO EXHBIITS 1050-1052" and "NI OBJECTIONS TO THE ADMISSIBILITY OF RAUCH EXHIBITS 1050, 1051 and 1052," respectively.

123. Ni has not provided evidence that it timely objected to exhibits RX 1025 and RX 1038.

124. Ni has not identified where in the record exhibits RX 1025, RX 1038, RX 1051, RX 1052, RX 1074 and RX 1075 were relied upon by its opponent Rauch.

125. Rauch's exhibit list filed when it submitted its record for decision on motion ("Rauch's exhibit list," Paper 87, p. 7) identifies exhibits RX 1052 and 1054 as documents upon which Rauch will rely to prove its earliest corroborated conception of the subject matter of the count in related interference 105,381.

126. Similarly, Rauch's exhibit list identifies exhibit RX 1074 as the declaration of Dr. Gavin R. Scranton filed in related interference 105,240 (Paper 87, p. 10).

127. Moreover, a cursory review of Rauch's exhibit list clearly indicates that exhibits RX 1025, RX 1038 and RX 1075 are not of record in this interference and, therefore, are not at issue in this interference (Paper 87, pp. 4-5).

128. Furthermore, according to Ni, "even though Party Rauch does not appear to have yet relied on any of Rauch Exhibits **1051** and **1052** in support of a motion. Party Ni serves Party Rauch with these objections to provide notice to Party Rauch that if, and when, any of Rauch Exhibits **1051** and **1052** are relied upon, the following objections will be raised, unless cured by Party Rauch" (NX 2193, p. 1, ¶ 1).

Thus, Ni has failed to object timely to evidence it seeks to exclude (RX 1025 and RX 1038). Furthermore, Ni is seeking to exclude evidence which is either not of record in this interference (RX 1025, RX 1038, RX 1074 and RX 1075) and/or has not been relied upon by Rauch in this interference (RX 1025, RX 1038, RX 1052, RX 1054, RX 1074 and RX 1075). Therefore, Ni miscellaneous motion 4 to exclude evidence is **denied** as to exhibits RX 1025, RX 1038, RX 1052, RX 1054, RX 1074, RX 1075 and NX 2179-2181.

129. Exhibit 1051 is identified in Rauch's exhibit list as a document said to prove Rauch's earliest corroborated conception of the invention of the count in this interference (Paper 87, p. 7).

According to 37 CFR § 41.204(a)(2)(iv), a party filing a priority statement must "[p]rovide a copy of the earliest document upon which the party will rely to show conception." Exhibit 1051 was submitted by Rauch in fulfillment of that requirement (FF 120). Ni does not contend that Rauch has relied on exhibit RX 1051 in support of any of Rauch's motion papers. The time for Rauch to lay a foundation for and authenticate its exhibit RX 1051 is when Rauch relies upon the exhibit, i.e., as part of its priority motion. The time for us to weigh the reliability and probative value of exhibit RX 1051 is when it is submitted into evidence as party of Rauch's priority motion when the motion is filed. Therefore, Ni miscellaneous motion 4 to exclude evidence is **denied** as to exhibit RX 1051.

130. As to the last evidence as issue, selected portions of the deposition testimony of Dr. Cheng (NX 2124, p. 132, l. 16 through p. 135, l. 5; p. 135, l. 9 through p. 136, l. 13 (with errata sheets)), Ni argues that this evidence was relied upon in Rauch reply 1 and Rauch reply 4 (Paper 80, p. B2, SMF 10).

131. Ni explicitly directs our attention (Paper 80, pp. 17-18) to the following testimony as an example of how the direct examination of Dr. Cheng violates FRE 611(c), FRE 403 and Cross Examination Guideline [3]:

> MR. WISE: Okay. Back on the record.
>
> Q. I want to have you focus on paragraph 10. Paragraph 10 you said, "The specification of the '861 application also contains additional substantial disclosure regarding antibodies to TRAIL-R, including methods for obtaining these antibodies and methods

of obtaining antigen binding fragments of these antibodies."

And it says "'861 application, page 13, line 14 to page 15, line 6."

Where in the specification of the '861 application would you find additional substantial disclosure relating to the antibodies for TRAIL-R?

A. You mean where I can find the information?

Q. Yes.

A. That's indicated here is the page 13 and the line 14 to 15, line 14 through page 15 of line 6.

Q. Okay. Can you direct me to that, please.

A. Where is the --

Q. You have that there. You were looking at the claims and you were going to show me support and specification.

MR. GOLDSTEIN: Objection.

THE WITNESS: So it's indeed in the page is 13, there is a title, "Antibodies" section, and talking about how antibody generated, including the monoclonal and polyclone antibodies.

MR. GOLDSTEIN: I am going to move to strike the question and the answer.

First, since Rauch responsive motion 4 was dismissed as moot, we did not reach Rauch reply 4. Second, Ni did not explain where and how Rauch relied upon the objected to portions of Dr. Cheng's testimony in Rauch reply 1 to support its position, e.g., how does Rauch rely upon this allegedly elicited testimony to support its motion 1 for benefit of the filing date of an earlier application for the subject matter of a count directed to isolated nucleotides. Third, to the extent Ni argues the objected portions of Dr. Cheng's testimony are

irrelevant, confusing or prejudicial, that objection goes to the weight to be accorded the testimony. We have accorded Dr. Cheng's testimony the weight appropriate to its relevance and the underlying facts and data relied upon in support of his opinion. Ni has not shown otherwise. Therefore, Ni miscellaneous motion 4 to exclude evidence is **denied** as to selected portions of the deposition testimony of Dr. Cheng (NX 2124, p. 132, l. 16 through p. 135, l. 5; p. 135, l. 9 through p. 136, l. 13 (with errata sheets)).

Therefore, Ni miscellaneous motion 4 is **denied**.

XI. Order

Based on the foregoing and for the reasons given, it is

ORDERED that Ni substantive motion 1 is **denied**;

FURTHER ORDERED that Rauch responsive motion 4 is **dismissed** as moot;

FURTHER ORDERED that Rauch substantive motion 1 is **granted**;

FURTHER ORDERED that Ni substantive motion 2 is **granted-in-part, denied-in-part** and **dismissed-in-part**;

FURTHER ORDERED that Rauch substantive motion 2 is **denied**;

FURTHER ORDERED that Rauch substantive motion 3 is **granted-in-part** to the extent Ni claims 287, 289-299, 351-361, 389-403, 431, 432, 434-442, 446, 448-458, 507-517, 581 and 623-632 are unpatentable under 35 U.S.C. § 102(e) as clearly anticipated by U.S. Patent 6,072,047;

FURTHER ORDERED that Rauch miscellaneous motion 5 is **dismissed** as moot; and,

FURTHER ORDERED that Ni miscellaneous motion 4 is **denied**.

/Richard E. Schafer/
RICHARD E. SCHAFER
Administrative Patent Judge

/Adriene L. Hanlon/
ADRIENE LEPIANE HANLON
Administrative Patent Judge

BOARD OF PATENT APPEALS AND INTERFERENCES

/Carol A. Spiegel/
CAROL A. SPIEGEL
Administrative Patent Judge

Enc: MOLECULAR BIOLOGY AND BIOTECHNOLOGY, R. Meyers, ed., VCH Publishers, Inc., New York, NY (1995), p . 860

MOLECULAR CELL BIOLOGY, second edition, Darnell et al., W.H. Freeman and Company, New York, NY (1990), pp. 44-48.

MICROBIOLOGY: An Introduction, Tortora et al., The Benjamin/Cummings Publishing Company, Inc., Menlo Park, California, (1982) pp. 111-112, copy enclosed; MCB, pp. 55-65.

CLINICAL DIAGNOSIS AND MANAGEMENT BY LABORATORY METHODS, sixteenth edition, J. B. Henry ed., W.B. Saunders Company, Philadephia (1979), Vol. I, p. 992.

cc (Federal Express):

Attorney for **HUMAN GENOME SCIENCES, INC**:

Jorge A. Goldstein, Esq.
STERNE, KESSLER, GOLDSTEIN & FOX, PLLC
1100 New York Avenue, N.W.
Washington, DC 20005-3934
Tel: 202-371-2600
Fax: 202-371-2540

Attorney for **GENETECH**:

Oliver R. Ashe, Jr. Esq.
ASHE, P.C.
11440 Isaac Newton Square North, Suite 210
Reston, VA 20190
Tel: 703-860-1515
Fax : 703-860-1515

Counsel for **IMMUNEX CORPORATION**:

Michael J. Wise, Esq.
PERKINS COIE, LLP
1620 26th Street,
6th Floor, South Tower
Santa Monica, CA 90404-4013
Tel: 310-788-3210
Fax: 310-788-3399

Molecular Biology and Biotechnology

A Comprehensive Desk Reference

Edited by
Robert A. Meyers

Robert A. Meyers, Ph.D.
3715 Gleneagles Drive
Tarzana, CA 91356, USA

Management Supervised by:
Chernow Editorial Services, Inc.,
1133 Broadway, Suite 721, New York, NY, USA

Cover design by: G & H SOHO, Inc.

Cover art courtesy of Dr. Alexander Wlodawer from Figure 1 of his article,
"AIDS, Inhibitor Complexes of HIV-1 Protease in."
Art prepared by Dr. Jacek Lubkowski.

Library of Congress Cataloging-in-Publication Data

Molecular biology and biotechnology : a comprehensive desk reference /
Robert A. Meyers, editor.
p. cm.
Includes bibliographical references and index.
ISBN 1-56081-569-8 (alk. paper). — ISBN 1-56081-925-1 (pbk.: alk. paper)
1. Molecular biology—Encyclopedias. 2. Biotechnology—Encyclopedias.
I. Meyers, Robert A. (Robert Allen), 1936-
QH506.M66155 1995
574.8'8'03—dc20 95-9063
CIP

Printed in the United States of America

ISBN 1-56081-569-8 (hardcover)

Printing History:
10 9 8 7 6 5 4 3 2

ISBN 1-56081-925-1 (softcover)

Printing History:
10 9 8 7 6 5 4 3 2

Published jointly by:

VCH Publishers, Inc.	VCH Verlagsgesellschaft mbH	VCH Publishers (UK) Ltd.
220 East 23rd Street	P.O. Box 10 11 61	8 Wellington Court
New York, NY 10010	D-6940 Weinheim	Cambridge CB1 1HW
USA	Federal Republic of Germany	United Kingdom

Fax: (212) 481-0897
E-mail address: order@vch.com

Sequences of nucleic acids in DNA and RNA and of amino acids in proteins define the primary structure of these molecules. Sequence analysis is carried out using computer programs that implement algorithms to determine sequence properties and to compare sequences. Sequence comparison can indicate whether an RNA or protein molecule or region of DNA is already known (identity) or has some degree of similarity to a known sequence. Sequence similarity may indicate similar structure or function. Sequence analysis can suggest the function of an unknown sequence based on the features it contains. Sequence analysis is a necessary preliminary to detailed experimental studies of structure, function, and interactions of biological macromolecules. Sequences are the information repository of the cell and a natural index to our growing understanding of cellular processes as dynamic systems of interactions between macromolecules.

1 PURPOSE OF SEQUENCE ANALYSIS

1.1 Prediction of Function

Sequences that are unlike any known sequence may still be made to yield information that can suggest their possible function. The function of nucleic acids and proteins depends on their structure and involves complex interactions in three dimensions. It is not presently understood whether it is possible, in general, to derive structure from sequence. Sequence alone is therefore often inadequate to determine function. Predictions made from sequence analysis need to be experimentally tested. Nevertheless, computer analysis of sequences is valuable in suggesting the most useful experiments to perform.

1.2 Revealing Similarity

The first thing to do with a newly determined sequence is to compare it with all known sequences. The outcome may show identity to a known sequence, which may prove disappointing if one is hoping for something new. Similarity to a known sequence may suggest something new that can be characterized with relatively little effort. A totally unknown sequence may be a frustrating result: considerable effort will be needed to understand its function.

Sequence comparison is a nontrivial pursuit, and both statistical and biological considerations are involved. Statistically significant similarities (under some model and at some chosen level of significance) may be biologically meaningless. Sequence motifs that are statistically nonsignificant in similarity may encode the same function (this is likely to occur because the statistical model based on sequence alone is incomplete). In an area fraught with such difficulties, common sense and interpretation based on utility are paramount.

Sequence dissimilarity can range from identity, difference due to sequencing errors, difference due to population polymorphism (individual variants), and differences in multiple copies of a gene in a single individual (multigene families) to wide evolutionary divergence of genes in different organisms. Sequences that are similar due to common function may not share a common ancestral sequence in biological evolution. In general, ideas about the evolutionary relationships of sequences are not experimentally testable. Sequence homology (similarity due to descent from a common ancestor) is a hypothesis, not an observable fact, except in the case of microbial populations with high mutation rates and short generation times, which may be studied experimentally through time.

2 ANALYSIS OF SINGLE SEQUENCES

2.1 DNA Composition, Isochores, and Codon Usage

Nucleotides in DNA sequences may be counted as singlets, doublets, or triplets in either strand. Doublets or triplets may be counted as overlapping or nonoverlapping in two or three phases, respectively, on either strand. The genomes of various organisms vary considerably in their DNA composition. Warm-blooded vertebrates have a higher G+C content, which correlates with the higher thermal stability of GC over AT base pairs. Composition of regions within a genome can also vary considerably. Mammalian genomes contain relatively GC-rich and AT-rich regions, which are called isochores. Overlapping doublet frequencies are highly characteristic for an organism. CG dinucleotides are less common than expected in vertebrates and angiosperms, probably because spontaneous deamination of 5-methylcytosine to thymine prevents the repair of methylated CpG. In DNA coding for protein, one phase of nonoverlapping triplets will be the phase of translation and the triplets will be codons. In a gene, the possible codons for each amino acid are unevenly used, and the frequency table for the 64 triplets is called codon usage. Codon usage is different between different species and between highly and lowly expressed gene in the same species.

2.2 Mapping DNA Sequence Features

Mapping the position of features on a DNA sequence is an important step in investigating its function. It is easy to map sites that can be precisely defined, such as stop codons or restriction enzyme recognition sites. Once DNA has been sequenced, the sizes of the fragments produced with any enzyme can be readily calculated. Features such as promoters, splice junctions, and ribosome binding sites are very difficult to predict because they are hard to specify. Mapping is most simply achieved by comparing the probe sequence with each position of the DNA sequence in turn and noting the hits. More sophisticated algorithms exist for rapid searching in large problems.

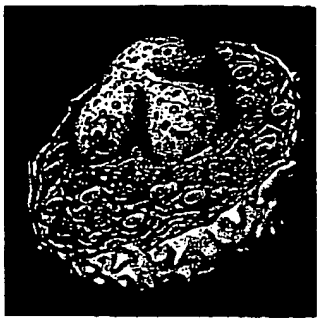
2.3 Repetitive Sequences

Direct repeats and inverted repeats (sometimes called dyad symmetries) are common in DNA from many sources. Mammalian genomes contain families of long (LINE) and short (SINE) repeats. Repeats of *L1* (*Kpn* I) type are 5000 to 7000 bp long and are present in the genome in 10^3 to 10^4 copies. Repeats of *Alu* type are 350 bp long and occur in as many as 9×10^4 copies. *Alu* repeats make human DNA hard to assemble from gel sequencing reads into the finished sequence. Inverted repeats occur in DNA coding for structural RNA, and these symmetry properties enable the RNA to fold into its secondary structure.

The dot plot is a diagram that reveals the presence of repeats and inverted repeats in sequences. It is also useful for comparing two different nucleic acid or protein sequences to detect regions of similarity. The dot plot is a rectangular array with rows labeled by one sequence and columns labeled by the other. A cell i, j can be used to represent the result of comparison of the jth residue of sequence A with the ith residue of sequence B. The simplest form of dot plot results from placing a diagonal mark in each cell where

Molecular Cell Biology

SECOND EDITION



JAMES DARNELL
Vincent Astor Professor
Rockefeller University

HARVEY LODISH
Member of the Whitehead Institute for Biomedical Research
Professor of Biology, Massachusetts Institute of Technology

DAVID BALTIMORE
President
Rockefeller University

SCIENTIFIC AMERICAN BOOKS

Distributed by W. H. Freeman and Company, New York

Cover illustration by Tomo Narashima

Library of Congress Cataloging-in-Publication Data

Darnell, James E.
Molecular cell biology / James Darnell, Harvey Lodish, David Baltimore.—2d ed.
p. cm.
Includes bibliographical references.
ISBN 0-7167-1981-9:—ISBN 0-7167-2078-7 (international student ed.):
1. Cytology. 2. Molecular biology. I. Lodish, Harvey F. II. Baltimore, David. III. Title.
[DNLM: 1. Cells. 2. Molecular Biology. QH 581.2 D223m]
QH581.2.D37 1990
574.87'6042—dc20
DNLM/DLC 89-70096
for Library of Congress CIP

Printed in the United States of America

Scientific American Books is a subsidiary of Scientific American, Inc.
Distributed by W. H. Freeman and Company, 41 Madison Avenue, New York, New York 10010 and 20 Beaumont Street, Oxford OX1 2NQ England

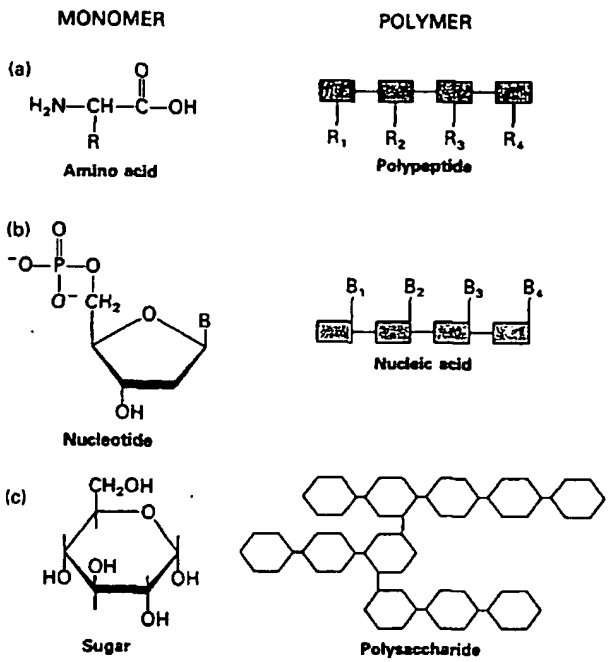
5 6 7 8 9 0 KP 9 9 8 7 6 5 4 3

ent amino acids in proteins. Thus a 100-unit protein has 20^{100} (more than 10^{130}) possible structures. This enormous variability means that cells and organisms can differ greatly in structure and function even though they are constructed of the same types of biopolymers produced by similar chemical reactions.

Starch (a storage form of glucose in plant cells), cellulose (a constituent of plant cell walls), and glycogen (a storage form of glucose in liver and muscle cells) are examples of another important type of biopolymer: the polysaccharide, *which is built of sugar monomers (Figure 2-1). At least 15 different monomeric sugars can be bonded in multiple ways to form various polysaccharides; thus many polysaccharides are nonlinear, branched molecules.*

Monomers are not the only small molecules important to cell structure. The lipids, for example, form the basic structure of cell membranes. Lipids cohere noncovalently in very large sheetlike complexes; the membranes thus formed are as crucial to living systems as are the biopolymers.

This chapter deals with the structures and some functions of biopolymers and small molecules; later chapters describe how the polymers are made and consider many of their other functions and interactions. ▲



▲ **Figure 2-1** (a) Proteins, linear biopolymers called polypeptides, are formed from monomeric subunits termed amino acids. Each of the 20 different amino acids has a different R group, or side chain. Thus the polypeptide shown here, which is constructed of four amino acids, has 20^4, or 160,000, possible structures. (b) Nucleic acids, also linear biopolymers, are formed from four monomers termed nucleotides, each of which has a different nitrogen-containing base structure (B). The nucleic acid shown here has 4^4, or 256, possible structures. (c) Polysaccharides are built of monomeric saccharide (sugar) subunits. Because sugar residues can bind to one another at different positions, nonlinear branching polymers are often formed. The rings in (b) and (c) are depicted as Haworth projections (planar structures with a hint of perspective).

Proteins

Proteins are the working molecules of the cell. They catalyze an extraordinary range of chemical reactions, provide structural rigidity, control membrane permeability, regulate the concentrations of metabolites, recognize and noncovalently bind other biomolecules, cause motion, and control gene function. These incredibly diverse tasks are performed by molecules constructed from only 20 different amino acids.

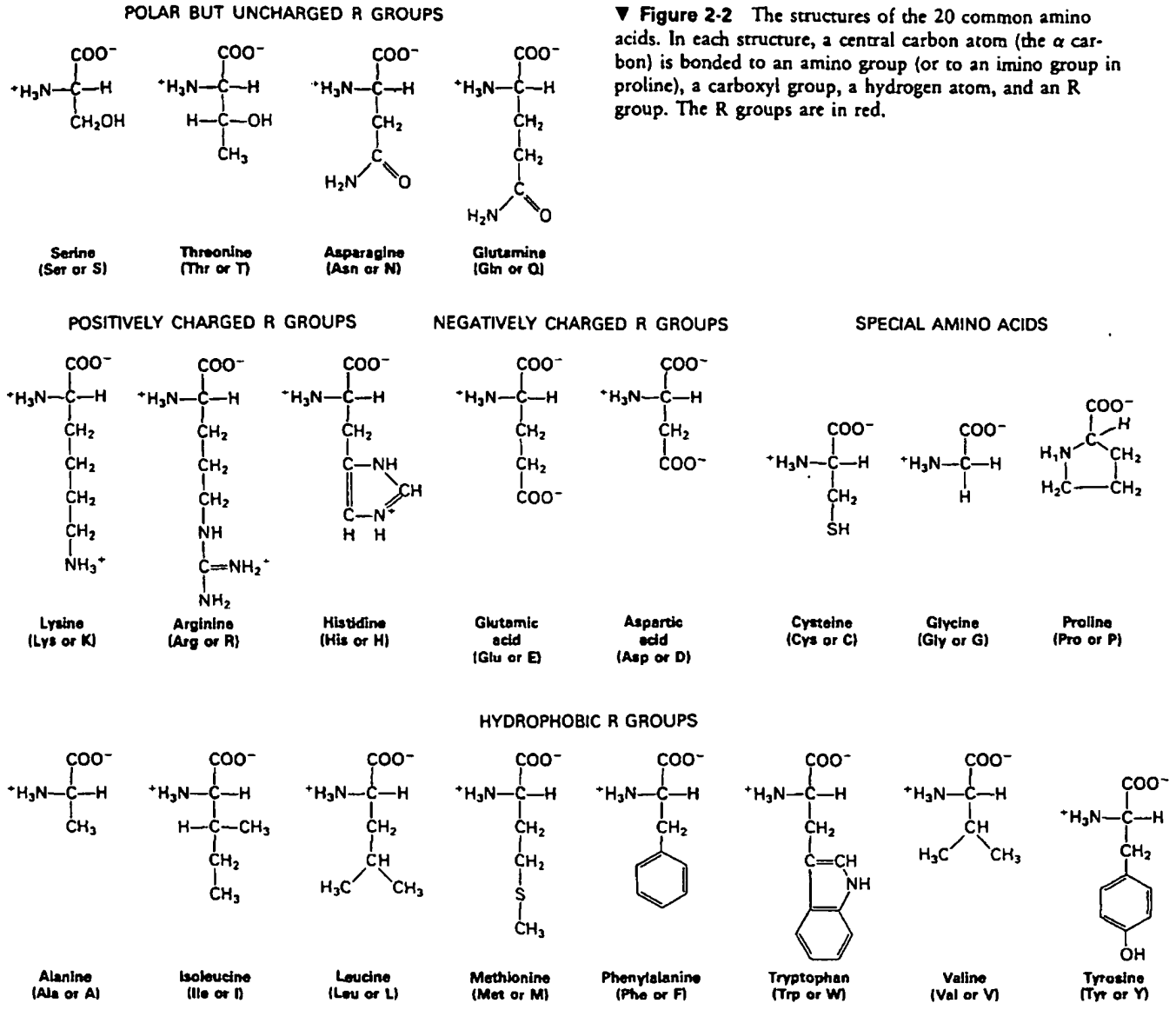
Amino Acids—the Building Blocks of Proteins—Differ Only in Their Side Chains

The monomers that make up proteins are called amino acids because, with one exception, each contains an *amino group* ($—NH_2$) and an acidic carboxyl group (—COOH). The exception, proline, has an *imino group* (—NH—) instead of an amino group. At typical pH values in cells, the amino and carboxyl groups are ionized as $—NH_3^+$ and $—COO^-$. All amino acids are constructed according to a basic design: a central carbon atom, called the α carbon C_α (because it is adjacent to the acidic carboxyl group), is bonded to an amino (or imino) group, to the carboxyl group, to a hydrogen atom, and to one variable group, called a *side chain* or R *group* (Figure 2-2). The side chains give the amino acids their individuality.

The amino acids represent the alphabet in which linear proteins are "written"; any student of biology must be familiar with the special properties of each letter of this alphabet. These letters can be classified into a few distinct categories.

The side chains of four of the amino acids are highly ionized and therefore charged at neutral pH. *Arginine* and *lysine* are positively charged; *aspartic acid* and *glutamic acid* are negatively charged and exist as asparate and glutamate. The side chain of a fifth amino acid, *histidine*, is positively charged, but only weakly at neutral pH. In many cases, arginine may substitute for lysine, or aspartate for glutamate, with little effect on the structure or function of the protein.

Serine and *threonine*, whose side chains have an —OH group, can interact strongly with water by forming hydrogen bonds. The side chains of *asparagine* and *gluta-*



▼ **Figure 2-2** The structures of the 20 common amino acids. In each structure, a central carbon atom (the α carbon) is bonded to an amino group (or to an imino group in proline), a carboxyl group, a hydrogen atom, and an R group. The R groups are in red.

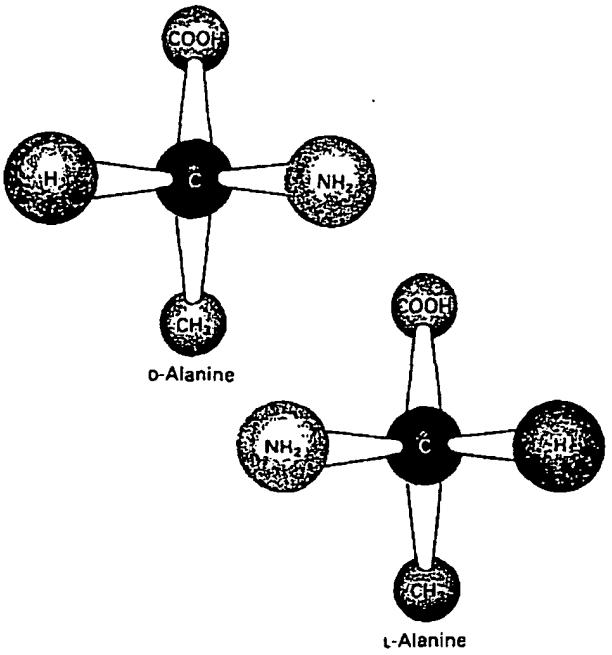
mine have polar amide groups with even more extensive hydrogen-bonding capacities. Together with the charged amino acids, these amino acids constitute the nine hydrophilic or polar amino acids.

The side chains of several other amino acids—*alanine, isoleucine, leucine, methionine, phenylalanine, tryptophan,* and *valine*—consist only of hydrocarbons, except for the sulfur atom in methionine and the nitrogen atom in tryptophan. These nonpolar amino acids are hydrophobic; their side chains are only slightly soluble in water. *Tyrosine* is also strongly hydrophobic because of its benzene ring, but its hydroxyl group allows it to interact with water, making its properties somewhat ambiguous.

Cysteine plays a special role in proteins because its —SH group allows it to dimerize through an —S—S— bond to a second cysteine, thus covalently linking regions of polypeptide to one another. When the —SH remains free, cysteine is quite hydrophobic.

Two other special amino acids are *glycine* and *proline*. Glycine has a hydrogen atom as its R group; thus it is the smallest amino acid and has no special hydrophobic or hydrophilic character. Proline, as an imino acid, is very rigid and creates a fixed kink in a polypeptide chain. It is quite hydrophobic.

The structure of all amino acids except glycine are asymmetrically arranged around the α carbon, because it is bonded to four different atoms or groups of atoms

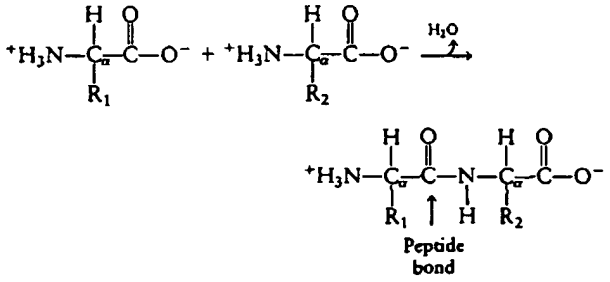


▲ **Figure 2-3** Stereoisomers of the amino acid alanine. The α carbon is black.

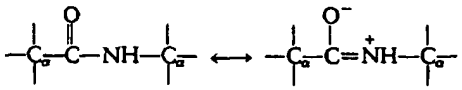
(—NH_2, —COOH, —H, and —R). Thus all amino acids except glycine can have one of two stereoisomeric forms. By convention, these mirror-image structures are called the D and the L forms of the amino acid (Figure 2-3). They cannot be interconverted without breaking a chemical bond. With rare exceptions, only the L forms of amino acids are found in proteins.

Polypeptides Are Polymers Composed of Amino Acids Connected by Peptide Bonds

The *peptide bond,* the chemical bond that connects two amino acids in a polymer, is formed between the amino group of one amino acid and the carboxyl group of another. This reaction, called *condensation,* liberates a water molecule:

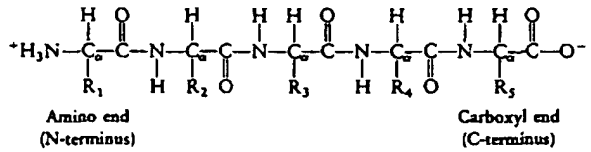


Because the carboxyl carbon and oxygen atoms are connected by a double bond, the peptide bond between carbon and nitrogen exhibits a partial double-bond character, as shown by the resonance structures



making it shorter than the typical C—N single bond. The six atoms of the peptide group (the two carbons of the adjacent amino acids and the carbon, oxygen, nitrogen, and hydrogen atoms of the bond) lie in the same plane (Figure 2-4a). However, adjacent peptide groups are not necessarily coplanar, due to rotation about the C—C_α and N—C_α bonds (Figure 2-4b).

A single linear array of amino acids connected by peptide bonds is called a *polypeptide.* If the polypeptide is short (fewer than 30 amino acids long), it may be called an *oligopeptide* or just a *peptide.* Polypeptides in living cells differ greatly in length; they generally contain between 40 and 1000 amino acids. Each polypeptide has a free amino group at one end (the N-terminus) and a free carboxyl group at the other (the C-terminus):

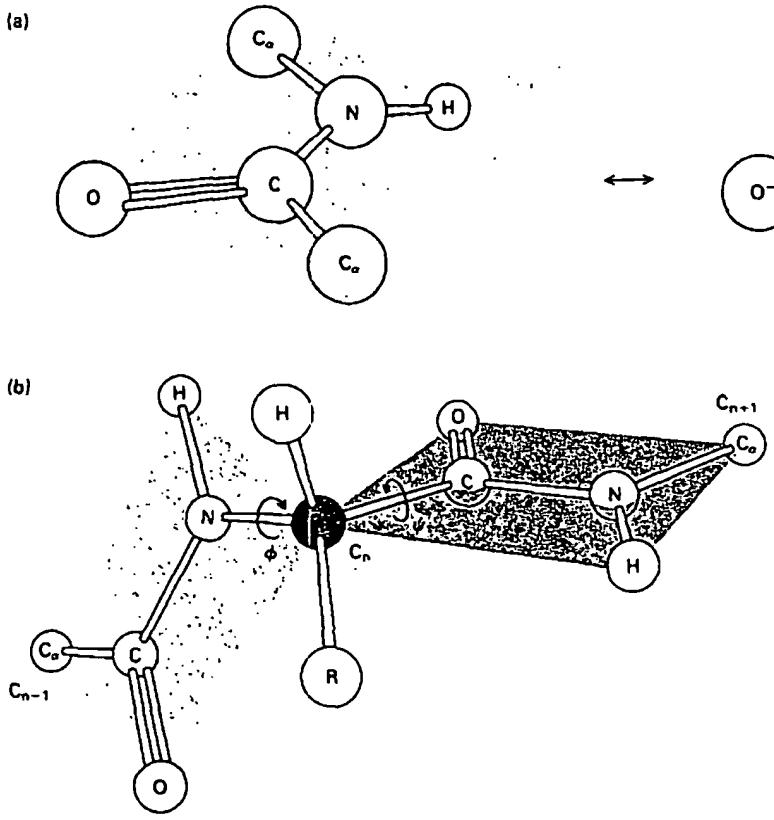


A protein is not merely a linear string of amino acids. The polypeptide folds up to form a specific three-dimensional structure that can be a long rod, as in the *fibrous proteins* that give tissues their rigidity, or a compact ball called a *globular protein,* as in many proteins that catalyze chemical reactions (enzymes), or a combination of balls and rods. The polypeptide can be modified further by the covalent or noncovalent attachment of additional small molecules.

A protein adopts a stable, folded conformation mainly through noncovalent (ionic, hydrogen, van der Waals, and hydrophobic) interactions. Its stability is also enhanced by the formation of covalent disulfide bonds between cysteines in different parts of the chain. Proteins may also consist of multiple polypeptide chains held together by noncovalent forces and, in some cases, by disulfide bonds. A well-characterized example is the hemoglobin molecule, which consists of four chains: two identical α chains and two identical β chains (Figure 2-5).

Three-dimensional Protein Structure Is Determined through X-ray Crystallography

The detailed three-dimensional structures of numerous proteins have been established by the painstaking efforts of many workers—notably, Max Perutz and John Kendrew, who perfected the x-ray crystallography of



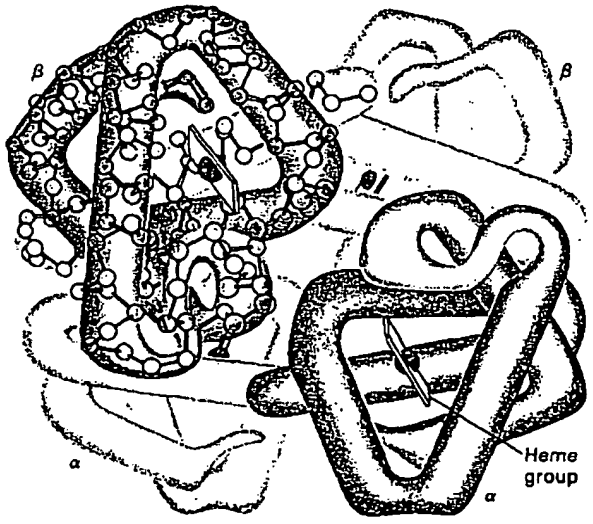
◀ **Figure 2-4** (a) Because the carbon-nitrogen peptide bond has a partial double-bond character, the peptide group is planar. (b) However, there is considerable flexibility in the geometry of polypeptides: rotation is possible about the two covalent single bonds that connect each α carbon to the two adjacent planar peptide units. But some restrictions do apply to the values of ψ and ϕ. For example, if the pictured adjacent peptide groups were coplanar, then certain oxygen and hydrogen atoms would be separated by less than their van der Waals radii and would repel one another.

proteins, in which beams of x-rays are passed through a crystal of protein. The wavelengths of x-rays are about 0.1–0.2 nanometers (nm)—short enough to resolve the atoms in the protein crystal. The three-dimensional structure of the protein can be deduced from the *diffraction pattern* of discrete spots that is produced when the scattered radiation is intercepted by photographic film. Such patterns are extremely complex; as many as 25,000 diffraction spots can be obtained from a small protein. Elaborate calculations and modifications of the protein (such as binding of heavy metal) must be made to interpret the diffraction pattern and to solve the structure of the protein.

Recently, three-dimensional structures of some small proteins have been determined by nuclear magnetic resonance (nmr) methods. An advantage of this approach is that it avoids the need to crystallize the protein. A disadvantage is that it is limited to relatively small proteins (up to about 20,000 molecular weight).

The Structure of a Polypeptide Can Be Described at Four Levels

The structures adopted by polypeptides can be divided into four levels of organization. *Primary structure* refers to the linear arrangement of amino acid residues along a



▲ **Figure 2-5** The conformations assumed by the two α and two β chains in a molecule of hemoglobin. Each chain forms several α helices (see Figure 2-6). Only the backbones formed by the carbon and nitrogen atoms of the chains are shown here. A multitude of noncovalent interactions stabilize the conformations of the individual chains and the contacts between them. A heme group is bound to each chain. *After R. E. Dickerson and I. Geis, 1969,* The Structure and Action of Proteins, *Benjamin-Cummings, p. 56. Copyright 1969 by Irving Geis.*

polypeptide chain and to the locations of covalent bonds (mainly —S—S— bonds) between chains. *Secondary structure* pertains to the folding of parts of these chains into regular structures, such as α helices and β pleated sheets. *Tertiary structure* includes the folding of regions between α helices and β pleated sheets, as well as the combination of these secondary features into compact shapes (domains). *Quaternary structure* refers to the organization of several polypeptide chains into a single protein molecule, such as in hemoglobin.

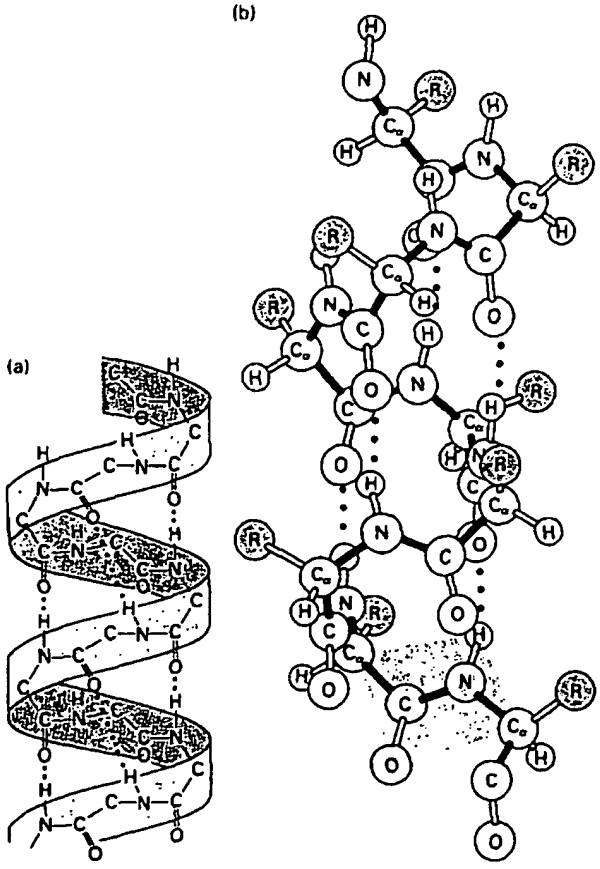
Two Regular Secondary Structures Are Particularly Important

The α Helix Although some regions of proteins are held in unique and irregular conformations, much protein structure involves repeated use of a limited number of regular configurations. One common structure, the *α helix*, was first described by Linus Pauling and Robert B. Corey in 1951. Through careful model building, these scientists came to realize that polypeptide segments composed of certain amino acids tend to arrange themselves in regular helical conformations. In an α helix, the carboxyl oxygen of each peptide bond is hydrogen-bonded to the hydrogen on the amino group of the fourth amino acid away (Figure 2-6), so that the helix has 3.6 amino acids per turn. Each amino acid residue represents an advance of about 1.5 Å along the axis of the helix. Every C=O and N—H group in the peptide bonds participates in a hydrogen bond, and the rigid planarity of the peptide bonds contributes to the rigid shape of the helix. In this inflexible, stable arrangement of amino acids, the side chains are positioned along the outside of a cylinder. The hydrogen-bonding potential of the peptide bonds is entirely satisfied internally, so that the polar or nonpolar quality of the cylindrical surface is determined entirely by the side chains. At least some of the amino acids in most proteins are organized into α helices.

Certain amino acid sequences adopt the α-helical conformation more readily than others. What determines this propensity is complicated, but some simple factors are evident. For instance, proline is rarely found in α-helical regions because it cannot use its peptide nitrogen to make a hydrogen bond. Glycine also is an infrequent participant. Another inhibiting factor can be the tendency of multiple identically charged residues to repel each other.

The α helix is a rodlike element of protein structure that serves many functions. A globular protein can be made up of short α-helical rods connected by bends that allow the rods to interact with each other; hemoglobin, for instance, is 70 percent α helical (see Figure 2-5). Alternatively, a single rod can span a long distance, as in the protein on the surface of the influenza virus (Figure 2-7a). Even in extended molecules, a,b,c the α helix is usually found packed against other elements of protein, not as an isolated structure. Long fibers, such as the skin protein keratin or the muscle protein myosin (Figure 2-7b), can be formed by two or three α helices that wrap gently around each other to form *coiled coils*. Small rods of α helix interact with DNA in some DNA-binding proteins (Figure 2-7c). A helical rod bearing only hydrophobic side chains can span lipid membranes well because the hydrophilic peptide bonds are buried inside the helix.

Many α helices are *amphipathic:* they expose hydrophilic side chains on one face and hydrophobic side chains on another face. Looking down the central axis of an α helix (Figure 2-8a), the amino acid residues are arranged in a wheel; if the helix is amphipathic, most or all

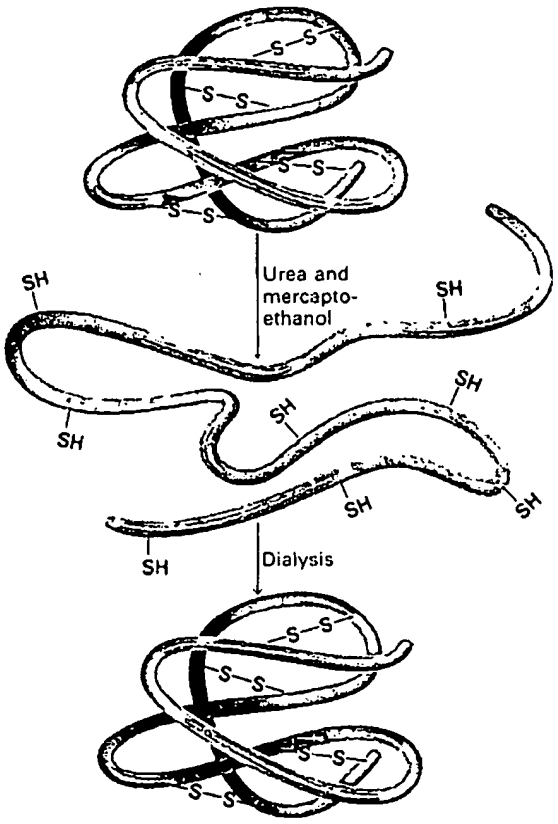


◀ **Figure 2-6** Models of the α helix. (a) This ribbonlike representation without R groups emphasizes the helical form. (b) This ball-and-stick representation emphasizes the role of the individual atoms and shows the R groups (green) that protrude from the helix body at regular intervals. Some of the planes of the C_α—CO—NH groups are shaded orange. *Part (b) after L. Stryer, 1988,* Biochemistry, *3d ed., W. H. Freeman and Company, p. 26.*

valently bound prosthetic group. For example, staphylococcal nuclease—a bacterial enzyme of 149 residues that degrades DNA and RNA—is totally denatured in acid but renatures to its native conformation within 0.1 s after the solution is neutralized. The three-dimensional architecture of this protein is solely a consequence of interactions among its amino acids and with its aqueous environment. In such cases, the genetic program of the cell must only define the primary structure of the protein—the amino acid sequence—and the tertiary structure is assured. With care, most proteins can be carried through a denaturation-renaturation cycle. Thus it is generally true that linear structure determines three-dimensional architecture.

The native form of some proteins is not the conformation with the lowest free energy and consequently cannot be completely restored on renaturation. This is particularly true of multichain proteins. The two chains of insulin, for example, can be separated by a combination of reducing agents (to break the disulfide bridges) and concentrated solutions of such chemicals as urea (to disrupt hydrogen and hydrophobic bonds). When the insulin renatures in the presence of oxidizing agents that promote the formation of disulfide bridges, a number of stable multichain aggregates do form, but *native* insulin molecules make up only a minor proportion of them. In the others, the re-formed disulfide bridges connect inappropriate parts of the chain.

Insulin is formed by the partial proteolysis (breaking down) of proinsulin, its larger precursor (see Figure 2-13). Denatured proinsulin, as opposed to the denatured two-chain form of insulin, can renature to form the native structure of proinsulin with a high efficiency. Presumably, within the cell, either proinsulin or preproinsulin folds in such a way that the correct disulfide bridges form at the lowest free energy. The cell utilizes these intermediate stages to form insulin, whose stable conformation is not the one of lowest free energy.



▲ **Figure 2-15** Denaturation and renaturation of a protein. Most polypeptides can be completely unfolded by treatment with an 8 *M* urea solution containing mercaptoethanol ($HSCH_2CH_2OH$). The urea breaks intramolecular hydrogen and hydrophobic bonds, and the mercaptoethanol reduces each disulfide bridge to two —SH groups. When these chemicals are removed by dialysis, the —SH groups on the unfolded chain oxidize spontaneously to re-form disulfide bridges, and the polypeptide chain simultaneously refolds into its native configuration.

Enzymes

Protein catalysts called *enzymes* are mediators of the dynamic events of life; almost every chemical reaction in a cell is catalyzed by an enzyme. Like other catalysts, enzymes increase the rates of reactions that are already energetically favorable; more precisely, enzymes increase the rates of forward and reverse reactions by the same factor. The name of an enzyme usually indicates its function: the suffix *-ase* is commonly appended to the name of the type of molecule on which the enzyme acts. Thus proteases degrade proteins, phosphatases remove phosphate residues, and ribonuclease cleaves RNA molecules.

The chemicals that undergo a change in a reaction catalyzed by an enzyme are the *substrates* of that enzyme. Because little free energy may be liberated in either direction in reversible reactions, the distinction between chemicals that are substrates and those that are products is often arbitrary.

Most enzymes are found inside cells, but a number are secreted by cells and function in the blood, the digestive tract, or other extracellular spaces. In microbial species, some enzymes function outside the organism. The number of different types of chemical reactions in any one cell is very large: an animal cell, for example, normally contains 1000–4000 different types of enzymes, each of which catalyzes a single chemical reaction or set of closely related reactions. Certain enzymes are found in the majority of cells because they catalyze common cellular reactions—the synthesis of proteins, nucleic acids, and phospholipids and the conversion of glucose and oxygen into carbon dioxide and water, which produces most of the chemical energy used in animal cells. Other enzymes are

found only in a particular type of cell within an organism, such as a liver cell or a nerve cell, because they carry out some chemical reaction unique to that cell. Also, many mature cells, including erythrocytes (red blood cells) and epidermal (skin) cells, may no longer be capable of making proteins or nucleic acids yet these cells still contain specific sets of enzymes that they synthesized at an earlier stage of differentiation.

Certain Amino Acids in Enzymes Bind Substrates: Others Catalyze Reactions on the Bound Substrates

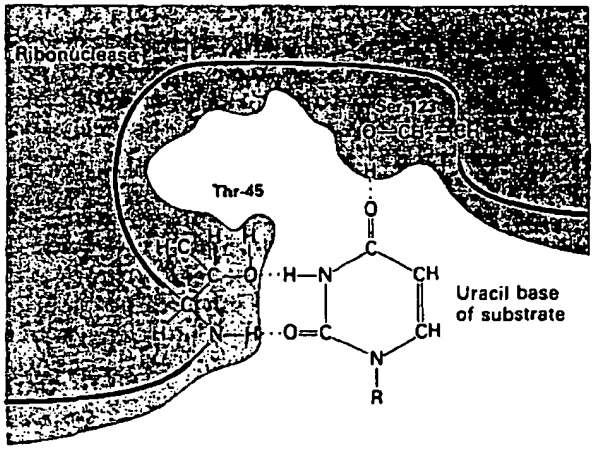
Two striking properties characterize all enzymes: their enormous *catalytic power* and their *specificity*. Quite often, the rate of an enzymatically catalyzed reaction is 10^6–10^{12} times that of an uncatalyzed reaction under otherwise similar conditions. The specificity of an enzyme is determined by the different rates at which it catalyzes closely similar chemical reactions or by its ability to distinguish between closely similar substrates.

Certain amino acid side chains of an enzyme are important in determining its specificity and its ability to accelerate the reaction rate. The properties of an enzyme are thus functions of its linear arrangement of amino acids and of the appropriate foldings of the peptide chain. Enzyme molecules have two important regions, or sites: one that recognizes and binds the substrate(s), and one that catalyzes the reaction once the substrate(s) have been bound. The amino acids in each of these key regions do not need to be adjacent in the linear polypeptide; they are brought into proximity in the folded molecule. In some enzymes, the catalytic site is part of the substrate-binding site. These two regions are called, collectively, the *active site*.

The binding of a substrate to an enzyme usually involves the formation of multiple noncovalent ionic, hydrogen, and hydrophobic bonds and van der Waals interactions (Figure 2-16). The array of chemical groups in the active site of the enzyme is precisely arranged so that the specific substrate can be more tightly bound than any other molecule (with the exception of some enzyme inhibitors) and the reaction can occur readily. In catalysis, covalent bonds between the enzyme and the substrate may be formed (and then broken) to reduce the activation energy for the reaction.

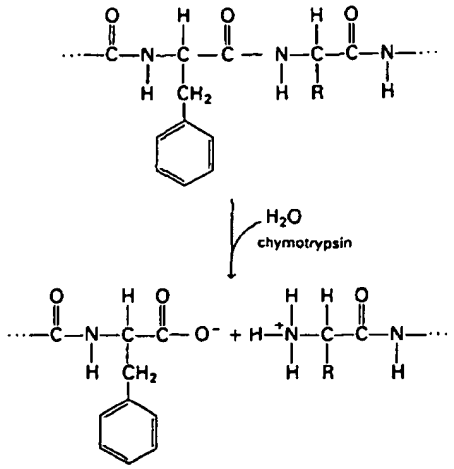
Trypsin and Chymotrypsin Are Well-characterized Proteolytic Enzymes

The proteolytic (protein-digesting) enzymes trypsin and chymotrypsin are synthesized in the pancreas and secreted into the small intestine as inactive precursors, or *zymogens*, called trypsinogen and chymotrypsinogen, respectively. These zymogens are not activated until they reach the small intestine where they hydrolyze peptide

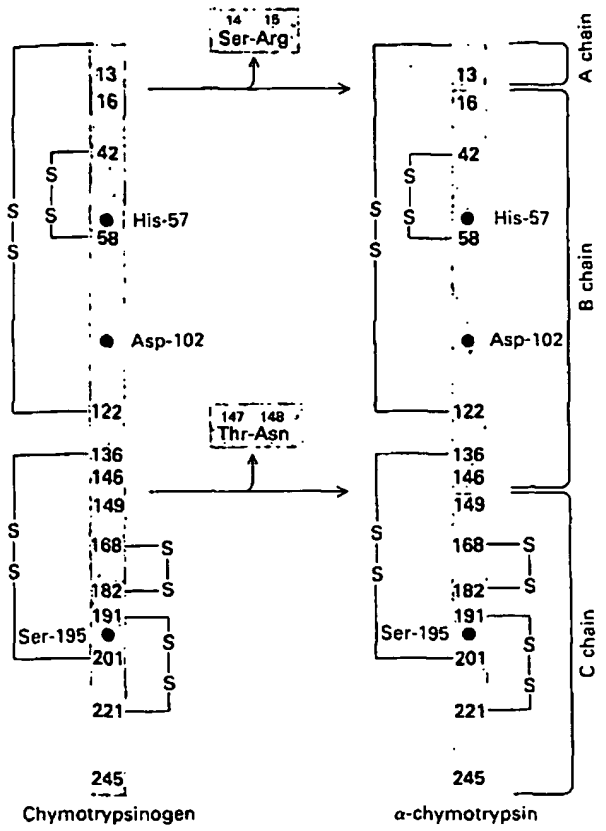


▲ **Figure 2-16** The specific binding of a substrate to an enzyme involves the formation of multiple noncovalent bonds. Here, two amino acid residues of the enzyme ribonuclease bind uracil, part of its substrate, by three hydrogen bonds. Substrates without the two C=O groups and one N—H group in the appropriate positions would be unable to bind or would bind less tightly. Other regions of the enzyme, not depicted here, bind other parts of the RNA substrate by hydrogen bonds and van der Waals interactions.

bonds of ingested proteins—a step in their digestion to single amino acids (Figure 2-17). The delay in activation serves an important regulatory purpose: it prevents the enzyme from digesting the pancreatic tissue in which it was made. Two irreversible proteolytic cleavages activate chymotrypsin. One cleavage removes serine 14 (the serine at position 14) and arginine 15 from chymotrypsinogen; the other removes threonine 147 and asparagine 148



▲ **Figure 2-17** The hydrolysis of a peptide bond by chymotrypsin.



▲ **Figure 2-18** A linear representation of the conversion of chymotrypsinogen into chymotrypsin by the excision of two dipeptides. The positions of the disulfide bridges are indicated. In the folded molecule, histidine 57, aspartate 102, and serine 195 are located in the active site.

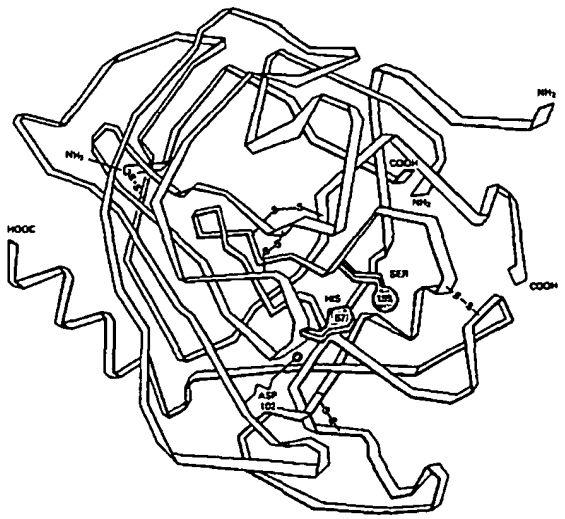
(Figure 2-18). Removal of these two dipeptides activates the protease function of the enzyme.

The hydrolysis of peptide bonds is energetically favorable ($\Delta G^{\circ\prime} = -2$ kcal/mol). Nonetheless, the activation energy for an *uncatalyzed* peptide-bond hydrolysis—say, in a neutral aqueous solution of a protein at room temperature—is so high that there is little or no hydrolysis even after several months. Biochemists can chemically hydrolyze proteins into their constituent amino acids by treating them with a 6 *M* solution of hydrochloric acid in an evacuated tube at 100°C for 24 h. Yet at 37°C and neutral pH, a molecule of trypsin or chymotrypsin can catalyze the hydrolysis of up to 100 peptide bonds per second. The power of enzymatically mediated catalysis is well-illustrated here: the addition of sufficient enzyme can do in seconds what otherwise would require harsh conditions and long times.

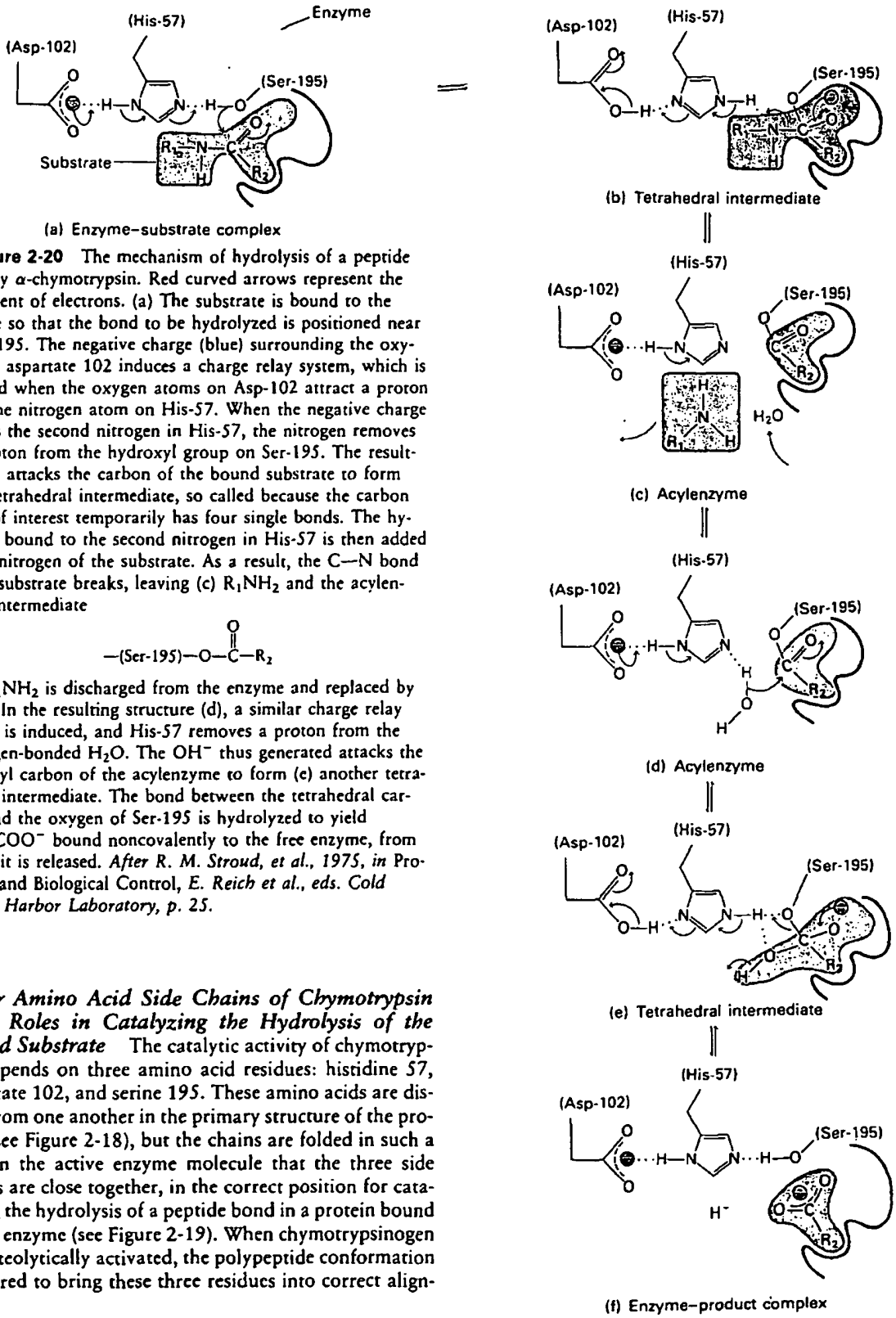
Chymotrypsin does not hydrolyze all peptide bonds; rather, it is selective for the peptide bond at the carboxyl ends of amino acids such as phenylalanine, tyrosine, and tryptophan, which have large hydrophobic side chains. Trypsin, by contrast, is specific for the peptide bond on the C-terminal side of lysine and arginine residues.

Specific Amino Acid Side Chains of Chymotrypsin Aid in Substrate Binding The reaction mechanism of chymotrypsin was deduced, in part, from the three-dimensional structure obtained by x-ray crystallography (Figure 2-19). The enzyme contains three polypeptides—the A, B, and C chains, which have 13, 131, and 97 amino acids, respectively. These chains are interconnected by disulfide bridges (see Figures 2-18 and 2-19). The molecule has two key structural features: the active site and the *hydrophobic cleft* (a crevice bordered by the side chains of several hydrophobic amino acid residues), which serves as the binding site for specific amino acid residues on the substrate. The conformation of this pocket allows the residues lining it to participate in hydrophobic interactions with the large hydrophobic side chains of phenylalanine, tyrosine, or tryptophan. Neither charged side chains nor small hydrophobic residues on the substrate can make the noncovalent bonds necessary to fit into the cleft.

The hydrophobic residues of most globular proteins are buried in the interior; when such proteins are in their native states, the peptide bonds linking the hydrophobic residues are not accessible to hydrolysis by chymotrypsin. Normally, stomach acids (pH 1) denature ingested proteins so that proteases in that organ can partly degrade them before their exposure to further digestion by chymotrypsin at neutral pH in the intestine.



▲ **Figure 2-19** A three-dimensional model of α-chymotrypsin determined from x-ray analysis. The N- and C-termini of the A, B, and C chains are indicated, as are the —S—S— bridges and the three amino acid residues of the active site (red). *After B. W. Matthews et al.*, 1967, Nature *214:652*.



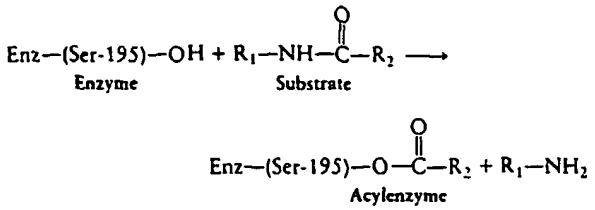
▲ **Figure 2-20** The mechanism of hydrolysis of a peptide bond by α-chymotrypsin. Red curved arrows represent the movement of electrons. (a) The substrate is bound to the enzyme so that the bond to be hydrolyzed is positioned near serine 195. The negative charge (blue) surrounding the oxygens in aspartate 102 induces a charge relay system, which is initiated when the oxygen atoms on Asp-102 attract a proton from the nitrogen atom on His-57. When the negative charge reaches the second nitrogen in His-57, the nitrogen removes the proton from the hydroxyl group on Ser-195. The resulting O^- attacks the carbon of the bound substrate to form (b) a tetrahedral intermediate, so called because the carbon atom of interest temporarily has four single bonds. The hydrogen bound to the second nitrogen in His-57 is then added to the nitrogen of the substrate. As a result, the C—N bond of the substrate breaks, leaving (c) R_1NH_2 and the acylenzyme intermediate

$$-(\text{Ser-195})-O-\overset{O}{\overset{\|}{C}}-R_2$$

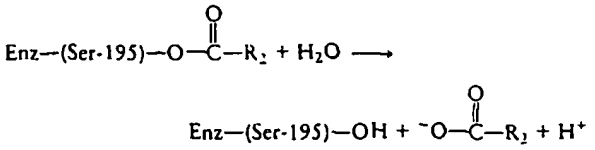
The R_1NH_2 is discharged from the enzyme and replaced by water. In the resulting structure (d), a similar charge relay system is induced, and His-57 removes a proton from the hydrogen-bonded H_2O. The OH^- thus generated attacks the carboxyl carbon of the acylenzyme to form (e) another tetrahedral intermediate. The bond between the tetrahedral carbon and the oxygen of Ser-195 is hydrolyzed to yield (f) R_2COO^- bound noncovalently to the free enzyme, from which it is released. *After R. M. Stroud, et al., 1975, in* Proteases and Biological Control, *E. Reich et al., eds. Cold Spring Harbor Laboratory, p. 25.*

Other Amino Acid Side Chains of Chymotrypsin Have Roles in Catalyzing the Hydrolysis of the Bound Substrate The catalytic activity of chymotrypsin depends on three amino acid residues: histidine 57, aspartate 102, and serine 195. These amino acids are distant from one another in the primary structure of the protein (see Figure 2-18), but the chains are folded in such a way in the active enzyme molecule that the three side chains are close together, in the correct position for catalyzing the hydrolysis of a peptide bond in a protein bound to the enzyme (see Figure 2-19). When chymotrypsinogen is proteolytically activated, the polypeptide conformation is altered to bring these three residues into correct alignment.

The hydrolysis reaction proceeds in two main steps. First, the peptide bond is broken and the carboxyl group is transferred to the hydroxyl residue of serine 195:



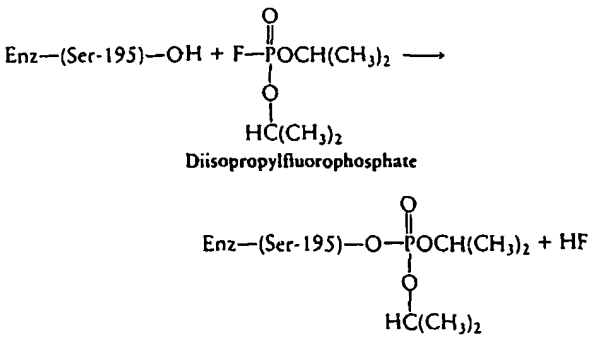
Second, this *acylenzyme* intermediate is hydrolyzed:



Note that the second step restores the enzyme to its original state.

Aspartate 102 and histidine 57 facilitate the acylation reaction by removing the proton from serine 195 and adding it to the nitrogen of the departing amino group (Figure 2-20). In a similar manner, aspartate 102 and histidine 57 facilitate the hydrolysis of the acylenzyme. These enzymatically catalyzed steps—transfer of a proton from the enzyme to the substrate, formation of a covalent acylserine intermediate, and hydrolysis of the acylenzyme—all drastically reduce the overall activation energy of the proteolysis reaction.

The hydroxyl group on serine 195 is unusually reactive. The concept of an "active" serine residue at the active site predated the determination of the crystal structure of chymotrypsin. It was already known, for example, that the compound diisopropylfluorophosphate is a potent inhibitor of chymotrypsin; it reacts only with the hydroxyl on serine 195 to form a stable covalent compound that irreversibly inactivates the enzyme:

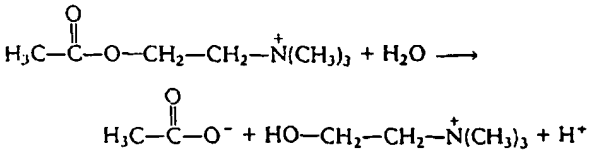


Trypsin and Chymotrypsin Have Different Substrate-binding Sites A comparison of trypsin and chymotrypsin will emphasize the nature of the specificity of enzymatically catalyzed reactions. About 40 percent of the amino acids in these two molecules are the same; in particular, the amino acid sequences in the vicinity of the key serine residue are identical:

$$\overset{195}{\text{—Gly—Asp—Ser—Gly—Gly—Pro}}$$

The three-dimensional structures and catalytic mechanisms of these two enzymes are also quite similar, indicating that they evolved from a common polypeptide. The major difference between trypsin and chymotrypsin is found in the side chains of the amino acids that line the substrate-binding site. The negatively charged amino acids in this area of the trypsin molecule facilitate the binding of only positively charged (lysine or arginine) residues, instead of hydrophobic ones.

Other Hydrolytic Enzymes Contain Active Serine Other, mostly unrelated, hydrolytic enzymes also contain an active serine residue that is essential for catalysis. For example, acetylcholinesterase catalyzes the hydrolysis of the neurotransmitter acetylcholine to acetate and choline:

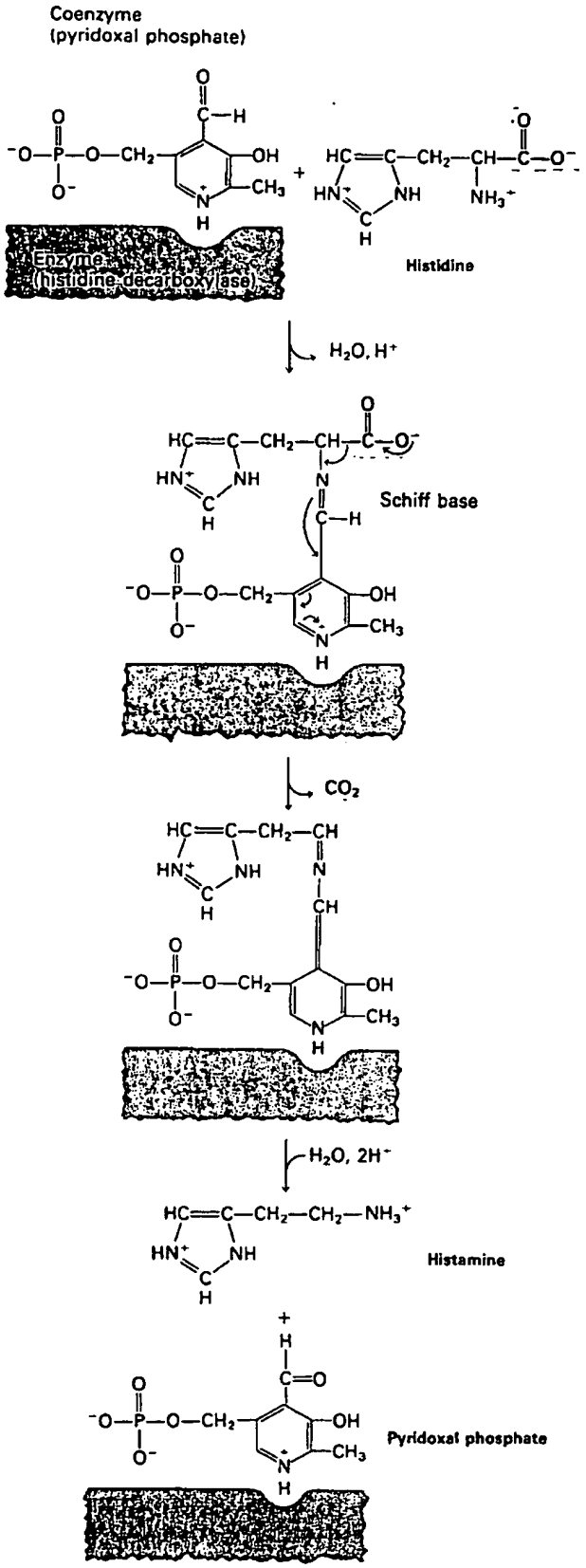


Diisopropylfluorophosphate is a potent, irreversible inhibitor of acetylcholinesterase as well as of chymotrypsin. The compound is lethal to animals because it blocks nerve transmission by causing a buildup of the transmitter substance. (The action of this transmitter is discussed in Chapter 20.)

Coenzymes Are Essential for Certain Enzymatically Catalyzed Reactions

Many enzymes contain a *coenzyme*—a tightly bound small molecule or prosthetic group essential to enzymatic activity. Vitamins required in trace amounts in the diet are often converted to coenzymes. Coenzyme A, for instance, is derived from the vitamin pantothenic acid; the coenzyme pyridoxal phosphate is derived from vitamin B_6. To cite just one example of how coenzymes function, we consider pyridoxal phosphate. The aldehyde group

can form a covalent complex called a *Schiff base* with an $—NH_2$ group of an amino acid, which facilitates or lowers the activation energy for the breaking of bonds to the carbon of the amino acid. Figure 2-21 shows how pyridoxal phosphate catalyzes the decarboxylation of histidine to form histamine—a potent dilator of small blood vessels. Histamine is released by certain cells in the course of allergenic hypersensitivity.



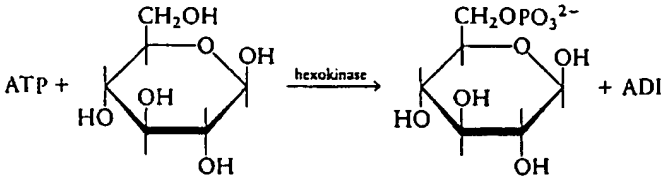
◀ **Figure 2-21** Pyridoxal phosphate, a coenzyme, participates in many reactions involving amino acids. When it is bound to histidine decarboxylase, as in this example, it forms a Schiff base with the α amino group of histidine. The positive charge on the nitrogen of pyridoxal phosphate then attracts the electrons from the carboxylate group of the histidine, via a charge relay system. This weakens the bond between the α carbon of the histidine and the carboxylate group, causing the release of CO_2. Finally, histamine, the reaction product, is hydrolyzed from the pyridoxal complex.

Substrate Binding May Induce a Conformational Change in the Enzyme

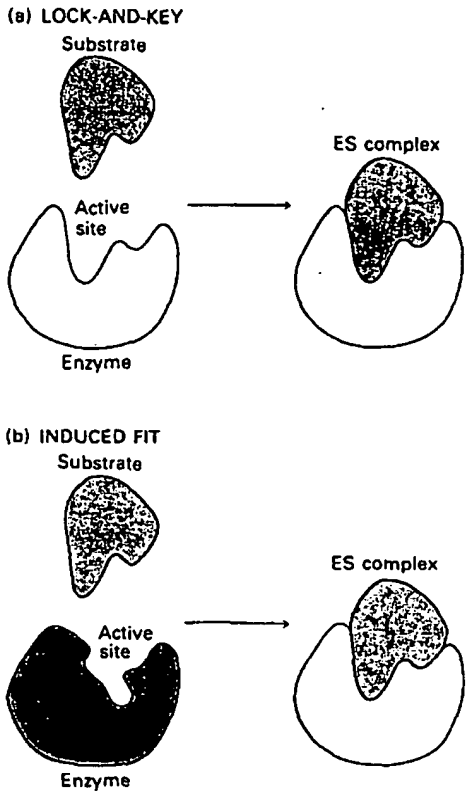
When a substrate binds to an enzyme, molecules of complementary charge or shape, or both, may simply fit together into a complex stabilized by a variety of noncovalent bonds. Such an interaction resembles the fitting of a key into a lock and is said to occur by a *lock-and-key* mechanism (Figure 2-22a).

In some enzymes, the binding of the substrate induces a conformational change in the enzyme that causes the catalytic residues to become positioned correctly. Molecules that attach to the substrate-binding site, or *recognition site*, of the enzyme but that do not induce a conformational change are not substrates of that enzyme. Thus an enzyme differentiates between a substrate and a nonsubstrate in two ways: Does the potential substrate bind to the enzyme? If so, does it induce the correct conformational change? When both criteria are met, the enzyme-substrate complex is said to demonstrate *induced fit* (Figure 2-22b).

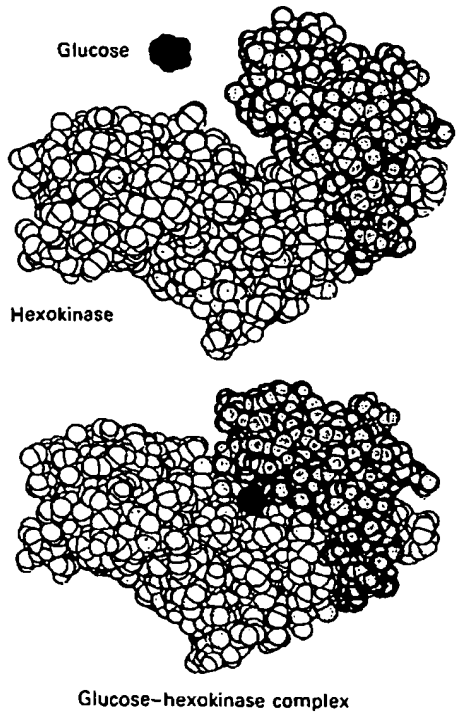
An important example of induced fit is provided by the enzyme hexokinase, which catalyzes the transfer of a phosphate residue from ATP to a specific carbon atom of glucose:



This is the first step in the degradation of glucose by cells. X-ray crystallography has shown that hexokinase consists of two domains. The binding of glucose induces a major conformational change that brings these domains closer together and creates a functional catalytic site (Figure 2-23). Only glucose and closely related molecules can induce this conformational change, ensuring that the enzyme is used to phosphorylate only the correct substrates. Molecules such as glycerol, ribose, and even water may bind to the enzyme at the recognition site but cannot induce the requisite conformational change, so they are not substrates for the enzyme.



▲ **Figure 2-22** Two mechanisms for the interaction of an enzyme and a substrate. (a) In the lock-and-key mechanism, the substrate fits directly into the binding site of the enzyme. (b) If binding occurs by induced fit, the substrate induces a conformational change in the enzyme that appropriately positions the substrate for catalysis.



▲ **Figure 2-23** The conformation of hexokinase changes markedly when it binds the substrate glucose: the two domains of the enzyme come closer together to surround the substrate. Molecules such as the five-carbon sugar ribose can also bind to hexokinase by forming specific hydrogen bonds with groups in the substrate-binding pocket of the enzyme, but only glucose can form all of the bonds that cause the enzyme to change its conformation. *Courtesy of Dr. Thomas A. Steitz.*

The Catalytic Activity of an Enzyme Can Be Characterized by a Few Numbers

Enzymatic specificity is usually quantified by discrimination ratios: a good substrate may be cleaved 10,000 times as fast as a poor substrate. The catalytic power of an enzyme on a given substrate involves two numbers: K_m, which measures the affinity of the enzyme for its substrate, and V_{max}, which measures the maximal velocity of enzymatic catalysis. Equations for K_m and V_{max} are most easily derived by considering the simple reaction

$$S \longrightarrow P$$
(substrate $\longrightarrow$ product)

in which the rate of product formation depends on [S], the concentration of the substrate, and on [E], the concentration of the catalytic enzyme. For an enzyme with a single catalytic site, Figure 2-24(a) shows how $d[P]/dt$, the rate of product production, depends on [S] when [E] is kept constant.

At low concentrations of S, the reaction rate is proportional to [S]; as [S] is increased the rate does not increase indefinitely in proportion to [S] but eventually reaches V_{max}, at which it becomes independent of [S]. V_{max} is proportional to [E] and to a catalytic constant k_{cat} that is an intrinsic property of the individual enzyme; halving [E] reduces the rate at all values of [S] by one-half.

When interpreting curves such as those in Figure 2-24, bear in mind that all enzymatically catalyzed reactions include at least three steps: (1) the binding of the substrate (S) to the enzyme (E) to form an enzyme-substrate complex (ES); (2) the conversion of ES to the enzyme-product complex (EP); and (3) the release of the product (P) from EP, to yield free P:

$$E + S \overset{\text{binding}}{\rightleftharpoons} ES \xrightarrow{\text{catalysis}} EP \xrightarrow{\text{release}} P + E$$

In the simplest case, the release of P is so rapid that we can write

$$E + S \underset{k_2}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_{cat}} E + P$$

The reaction rate $d[P]/dt$ is proportional to the concentration of ES and to the catalytic constant k_{cat} for the given enzyme:

$$\frac{d[P]}{dt} = k_{cat}\,[ES] \qquad (1)$$

To calculate [ES], we assume the reaction is in a steady state, so that $k_1\,[E]\,[S]$, the formation rate of [ES], is equal to the rate of its consumption, either by dissociation of uncatalyzed substrate at a rate of $k_2\,[ES]$ or by catalysis at a rate of $k_{cat}\,[ES]$:

$$k_1\,[E]\,[S] = (k_2 + k_{cat})\,[ES] \qquad (2)$$

If

$$[E]_{tot} = [E] + [ES] \qquad (3)$$

(where $[E]_{tot}$ is the sum of the free and the complexed enzyme, or the total amount of enzyme), then we can combine equations (2) and (3) to obtain

$$[E]_{tot} = [E] + [ES] = \frac{(k_2 + k_{cat})}{k_1[S]}\,[ES] + [ES]$$

$$= [ES]\left[1 + \left(\frac{k_2 + k_{cat}}{k_1}\right)\left(\frac{1}{[S]}\right)\right]$$

If we define K_m, called the *Michaelis* constant, as

$$\frac{k_2 + k_{cat}}{k_1} \qquad (4)$$

then

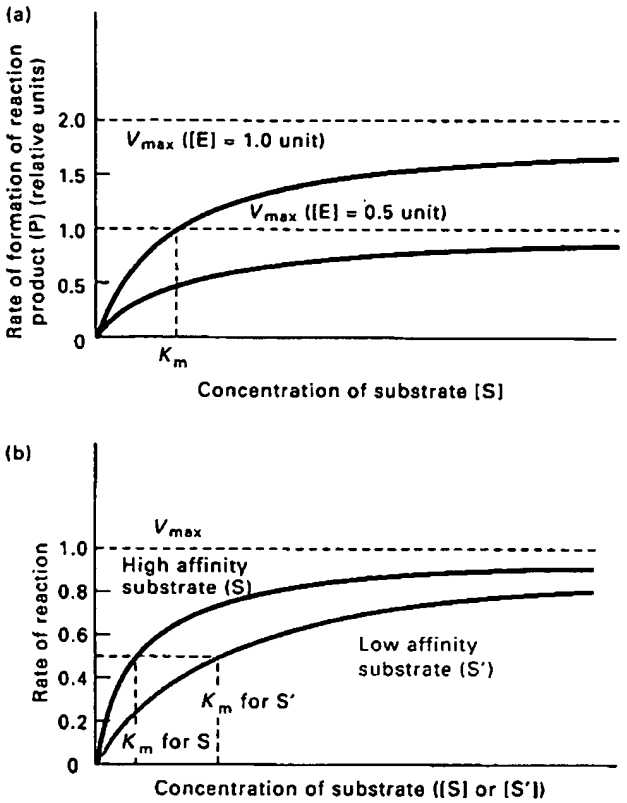
$$[ES] = \frac{[E]_{tot}}{1 + K_m/[S]}$$

Thus

$$\frac{d[P]}{dt} = k_{cat}\,[ES] = k_{cat}\,[E]_{tot}\,\frac{1}{1 + K_m/[S]}$$

$$= k_{cat}\,[E]_{tot}\,\frac{[S]}{[S] + K_m} \qquad (5)$$

This equation fits the curves shown in Figure 2-24a. V_{max}, which is equal to $k_{cat}\,[E]_{tot}$, is the maximal rate of product formation if all recognition sites on the enzyme are filled with substrate. K_m is equivalent to the substrate concentration at which the reaction rate is half-maximal. (If $[S] = K_m$, then from equation (5) we calculate the rate of product formation to be $\frac{1}{2}k_{cat}\,[E]_{tot} = \frac{1}{2}V_{max}$.) For most enzymes, the slowest step is the catalysis of [ES] to [E] + [P]. In these cases, k_{cat} is much less than k_2, so that $K_m = (k_2 + k_{cat})/k_1 \cong k_2/k_1$ is equal to the equilibrium constant for binding S to E. Thus the parameter K_m describes the affinity of an enzyme for its substrate. The smaller the value of K_m, the more avidly the enzyme can bind the substrate from a dilute solution (Figure 2-24b) and the lower the value of [S] needed to reach half-maximal velocity. The concentrations of the various small molecules in a cell vary widely, as do the K_m values for the different enzymes that act on them. Generally, the intracellular concentration of a substrate is approximately the same as or greater than the K_m value of the enzyme to which it binds.

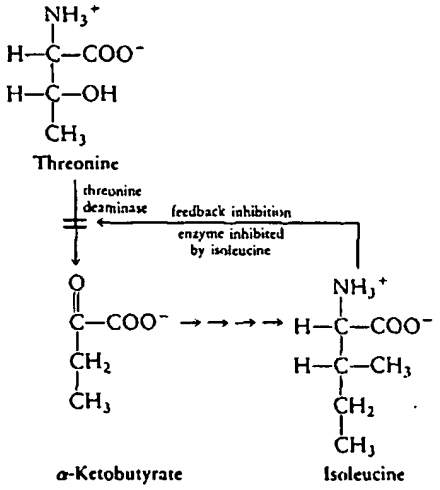


▲ **Figure 2-24** (a) The rate of a hypothetical enzymatically catalyzed reaction S → P for two different concentrations of enzyme [E] as a function of the concentration of substrate [S]. The substrate concentration that yields a half-maximal reaction rate is denoted by K_m. Doubling the amount of enzyme causes a proportional increase in the rate of the reaction, so that the maximal velocity V_{max} is doubled. The K_m, however, is unaltered. (b) The rates of reactions catalyzed by an enzyme with a substrate S, for which the enzyme has a high affinity, and with a substrate S', for which the enzyme has a low affinity. The V_{max} value is the same for S and S', but K_m is higher for S'.

The Actions of Most Enzymes Are Regulated

Many reactions in cells do not occur at a constant rate. Instead, the catalytic activity of the enzymes is *regulated* so that the amount of reaction product is just sufficient to meet the needs of the cell.

An Enzyme Can Be Feedback Inhibited in a Reaction Pathway Consider a series of reactions leading to the synthesis of the amino acid isoleucine, which is primarily used by cells as a monomer in the synthesis of proteins. The amount of isoleucine needed depends on the rate of protein synthesis in the cell. The first step in the synthesis of isoleucine is the elimination of an amino group, which converts the amino acid threonine to the compound α-ketobutyrate. Threonine deaminase—the enzyme that catalyzes this reaction—plays a key role in regulating the level of isoleucine. In addition to its substrate-binding sites for threonine, threonine deaminase contains a binding site for isoleucine. When isoleucine is bound there, the enzyme molecule undergoes a conformational change, so that it cannot function as efficiently. Thus isoleucine acts as an *inhibitor* of the reaction for the conversion of threonine. If the isoleucine concentration in the cell is high, the binding of isoleucine to the enzyme temporarily reduces the rate of isoleucine synthesis:



This is an example of *feedback inhibition*, whereby an enzyme that catalyzes one of a series of reactions is inhibited by the ultimate product of the pathway.

In isoleucine synthesis, as in most cases of feedback inhibition, the final product in the reaction pathway inhibits the enzyme that catalyzes the first step that does not also lead to other products. The suppression of enzyme function is not permanent. If the concentration of free isoleucine is lowered, bound isoleucine dissociates from the enzyme, which then reverts to its active conformation. The binding of the inhibitor isoleucine to the enzyme and its subsequent release can be described by the equilibrium-binding constant K_i, which is similar to the constant K_m used for substrate binding:

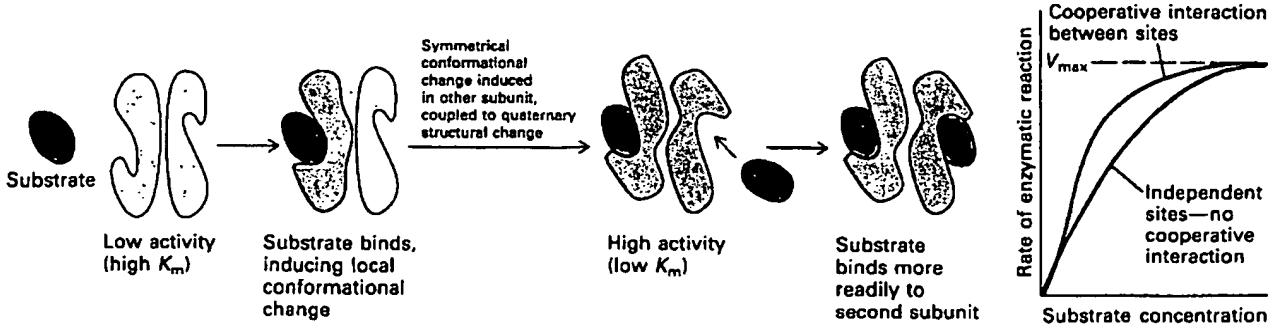
$$[E \cdot Ile]_{inactive} \overset{K_i}{\rightleftharpoons} [Ile] + [E]_{active}$$

$$K_i = \frac{[Ile][E]_{active}}{[E \cdot Ile]_{inactive}}$$

Many Enzymes Have Multiple Binding Sites for Regulatory Molecules Some enzymes have binding sites for small molecules that affect their catalytic activity; a stimulator molecule is called an *activator*. Enzymes may even have multiple sites for recognizing more than one activator or inhibitor. In a sense, enzymes are like microcomputers; they can detect concentrations of a variety of molecules and use that information to vary their own activities. Molecules that bind to enzymes and increase or decrease their activities are called *effectors*. Effectors can modify enzymatic activity because enzymes can assume both active and inactive conformations: activators are positive effectors; inhibitors are negative effectors. Effectors bind at *regulatory sites*, or *allosteric sites* (from the Greek for "another shape"), a term used to emphasize that the regulatory site is an element of the enzyme distinct from the catalytic site and to differentiate this form of regulation from competition between substrates and inhibitors at the catalytic site.

Multimeric Organization Permits Cooperative Interactions among Subunits Many enzymes and some other proteins are multimeric—that is, they contain several copies, or subunits, of one or more distinct polypeptide chains. Some multimeric enzymes contain identical subunits, each of which has a catalytic site and possibly an effector site. In other enzymes, regulatory sites and catalytic sites are located on different subunits, each with a particular structure. On binding an activator, inhibitor, or substrate, a subunit undergoes a conformational change, usually small, that triggers a change in quaternary structure. This quaternary rearrangement favors a similar conformational change in the other subunits, thereby increasing their affinity for the type of ligand initially bound (Figure 2-25). When several subunits interact cooperatively, a given increase or decrease in substrate or effector concentration causes a larger change in the rate of an enzymatic reaction than would occur if the subunits acted independently. Because of such *cooperative interactions*, a small change in the concentration of an effector or substrate can lead to large changes in catalytic activity.

Cooperative interactions among the four subunits in hemoglobin demonstrate clearly the advantages of multimeric organization. The binding of an O_2 molecule to any one of the four chains (each hemoglobin chain binds one O_2) induces a local conformational change in that subunit. This change can in turn induce a large change in quaternary structure. The quaternary change involves a rearrangement of the positions of the two α and two β chains in the tetramer. The local conformational changes that accompany O_2 binding can then occur more readily in the remaining subunits, increasing their affinity for oxygen. The binding of a second O_2 makes the quaternary structural change even more likely. The cooperative



▲ **Figure 2-25** A cooperative interaction between active sites (two identical subunits of a hypothetical enzyme). The binding of a substrate to one subunit of a multimeric enzyme induces a conformational change in the adjacent subunit, which lowers the K_m for the binding of the substrate there. Thus a small change in the substrate concentration can cause a much larger increase in the reaction rate than would occur if there were no cooperative interactions between active sites.

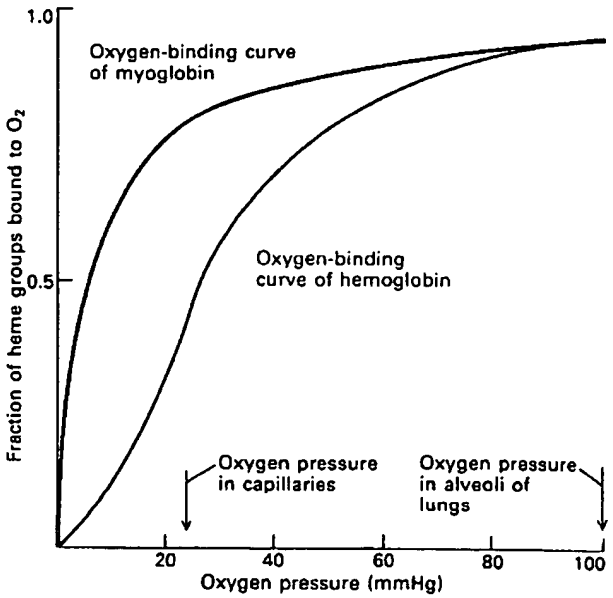
interaction between the chains causes the molecule to take up or lose four O_2 molecules over a much narrower range of oxygen pressures than it would otherwise. As a result, hemoglobin is almost completely oxygenated at the oxygen pressure in the lungs and largely deoxygenated at the oxygen pressure in the tissue capillaries (Figure 2-26).

The contrast between hemoglobin and myoglobin is revealing. Myoglobin is a single-chain oxygen-binding protein found in muscle. The oxygen-binding curve of myoglobin has the characteristics of a simple equilibrium reaction:

$$E + O_2 \overset{K_{O_2}}{\rightleftharpoons} E{-}K_{O_2}$$

Myoglobin has a greater binding affinity for O_2 (a lower K_{O_2}) than hemoglobin at all oxygen pressures. Thus, at the oxygen pressure in capillaries, O_2 moves from hemoglobin into the muscle cells, where it binds to myoglobin, ensuring the efficient transfer of O_2 from blood to tissues.

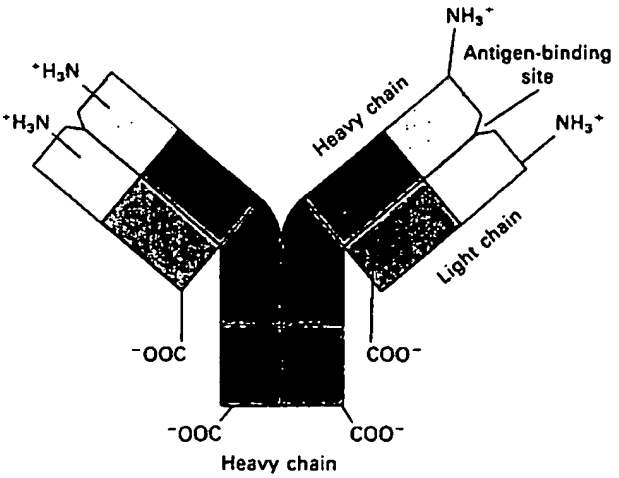
The quaternary-structure rearrangements associated with multimeric organization also provide a way for the effects of activator or inhibitor binding at an allosteric site to be transmitted to a distant catalytic site without large changes in the secondary or tertiary structure of an enzyme, which would be incompatible with the principle that a particular primary structure must adopt a unique folded conformation. Thus, for example, small conformational changes in a domain in response to binding of an effector molecule would produce a quaternary-structure change, which amplifies the conformational signal and allows it to be transmitted robustly to other parts of the enzyme, where it would induce a small conformational change affecting enzymatic activity. Membrane-embedded receptor proteins that must transmit a conformational signal from one side of a membrane to the other are also likely to be multimeric; they transmit the signal by quaternary-structure rearrangement or by an effector-induced shift in the monomer-multimer equilibrium.



◀ **Figure 2-26** The binding of oxygen to hemoglobin depends on cooperative interactions between the four chains. The graph shows the fraction of heme groups in hemoglobin and in myoglobin bound to O_2 as a function of the oxygen pressure. Note that the binding activity of hemoglobin increases sharply over a narrow range of oxygen pressures (20–40 mmHg). Hemoglobin is saturated with O_2 in the lungs, but it releases much of its bound O_2 at the low oxygen pressure in the tissue capillaries. At any oxygen pressure, myoglobin has a higher affinity for O_2 than hemoglobin does. As myoglobin is a principal muscle protein, this property allows oxygen to be transferred from blood to muscle.

Enzymes Are Regulated in Many Ways The activities of enzymes are extensively regulated so that the numerous enzymes in a cell work together harmoniously. All metabolic pathways are closely controlled at all times. Synthetic reactions occur when the products of these reactions are needed; degradative reactions occur when molecules must be broken down. Kinetic controls affecting the activities of key enzymes determine which pathways are going to be used and the rates at which they will function.

Regulation of cellular processes involves more than simply turning enzymes on and off, however. Some regulation is accomplished through *compartmentation*. Many enzymes are localized in specific compartments of the cell, such as the mitochondria or lysosomes, thereby restricting the substrates, effectors, and other enzymes with which an enzyme can interact. In addition, compartmentation permits reactions that might otherwise compete with one another in the same solution to occur simultaneously in different parts of a cell. Cellular processes are also regulated through the control of the rates of enzyme synthesis and destruction.



▲ **Figure 2-27** The structure of an antibody molecule illustrated in an immunoglobulin (IgG) made of four polypeptide chains: two identical heavy chains (blue) and two identical light chains (orange). Each antigen-binding site is formed by the N-terminal segments of a heavy and a light chain. The N-termini are highly variable in sequence, giving rise to the wide range of antibody specificity.

Antibodies

Enzymes are not the only proteins that bind tightly and specifically to smaller compounds. The insulin receptor on the surface of a liver cell, for example, can bind to insulin so tightly that the receptors on a cell are half-saturated when the insulin concentration is only 10^{-9} *M*. This protein does not bind to most other compounds present in blood; it mediates the specific actions of insulin on liver cells. A molecule other than an enzyme substrate that can bind specifically to a macromolecule is often called a *ligand* of that macromolecule.

The capacity of proteins to distinguish among different molecules is developed even more highly in blood proteins called *antibodies*, or *immunoglobulins*, than in enzymes. Animals produce antibodies in response to the invasion of an infectious agent, such as a bacterium or a virus. Antibodies will be discussed at length in Chapter 25. We introduce them here because they will appear as critical reagents in the discussions of many intervening chapters.

The recognition site of an antibody can bind tightly to very specific sites—generally on proteins or carbohydrates—on the surface of the infectious agent. Experimentally, animals produce antibodies in response to the injection of almost any foreign polymer; such antibodies bind specifically and tightly to the invading substance but, like enzymes, do not bind to dissimilar molecules. The antibody acts as a signal for the elimination of infectious agents. When it binds to a bacterium, virus, or virus-infected cell, certain white blood cells (leucocytes) recognize the invading body as foreign and respond by destroying it. The specificity of antibodies is exquisite: they can distinguish between proteins that differ by only a single amino acid and between the cells of different individual members of the same species.

All vertebrates can produce a large variety of antibodies, including ones that bind to chemically synthesized molecules. Exposure to an antibody-producing agent, called an *antigen*, causes an organism to make a large quantity of different antibody proteins, each of which may bind to a slightly different region of the antigen. For a given antigen, these constellations of antibodies may differ from one member of a species to another.

Antibodies are formed from two types of polypeptides: heavy chains, each of which is folded into four domains, and light chains, each of which is folded into two domains (Figure 2-27). The N-terminal domains of both heavy and light chains are highly variable in sequence, giving rise to the specific binding characteristics of antibodies.

Antibodies Can Distinguish among Closely Similar Molecules

The sequence of bovine insulin is identical to that of human insulin, except at three amino acids. Yet when bovine insulin is injected into people, some individuals respond by synthesizing antibodies that specifically recognize the specific amino acids in the bovine molecule, even though human beings generally do not produce anti-

MICROBIOLOGY

An Introduction

Gerard J. Tortora
Bergen Community College

Berdell R. Funke
North Dakota State University

Christine L. Case
Skyline College

The Benjamin/Cummings Publishing Company, Inc.
Menlo Park, California • Reading, Massachusetts
London • Amsterdam • Don Mills, Ontario • Sydney

Sponsoring Editor: James W. Behnke
Developmental Editor: Jane R. Gillen
Production Editor: Patricia S. Burner
Copy Editor: Janet Greenblatt
Book and Cover Designer: Michael A. Rogondino
Artist: Barbara Haynes
Illustrator: Michael Fornalski
Photo Researcher: Carl W. May

About the cover: A technician is isolating plasmids, which are tiny circles of DNA found in bacteria. The plasmids are dissolved in a dye solution that fluoresces pink under ultraviolet light. Genetic engineering using plasmids is revolutionizing the biological sciences and industry (see pages 226–229 and 704–707).

Figure acknowledgments begin on page 749.

Library of Congress Cataloging in Publication Data

Tortora, Gerard J.
Microbiology: an introduction.

Bibliography: p.
Includes index.
1. Microbiology. I. Funke, Berdell R.
II. Case, Christine L., 1948–
III. Title.
QR41.2.T67 576 81-21712
ISBN 0-8053-9310-2 AACR2

abcdefghij-DO-898765432

The Benjamin/Cummings Publishing Company, Inc.
2727 Sand Hill Road
Menlo Park, California 94025

tial energy and therefore serve as energy carriers to drive energy-requiring reactions. The most common energy carrier in all biological systems is adenosine triphosphate (ATP); its structure can be reviewed in Figure 2–20. The role of ATP in the relationship between catabolic and anabolic processes is shown in Figure 5–1.

A little later in the chapter, we will examine some representative chemical reactions that deal with energy production (catabolic reactions) and energy utilization (anabolic reactions) in microorganisms. We will then look at how these various reactions are integrated within the cell. But first let us consider the principal properties of a group of proteins involved in almost all biologically important chemical reactions. These proteins, the enzymes, were described briefly in Chapter 2. Although it is beyond the scope of this text to name and discuss the actions of individual enzymes, you should be aware of the central role of enzymes in metabolic reactions. It is important to understand that a cell's metabolic pathways are determined by its enzymes, which are, in turn, determined by its genetic makeup.

ENZYMES

Many organic chemicals are so stable that they could remain unchanged in a cell for years. To activate these chemicals, living cells produce **enzymes**, proteins that act as catalysts in chemical reactions of importance to the cell. A *catalyst* is a substance that speeds up a reaction without being

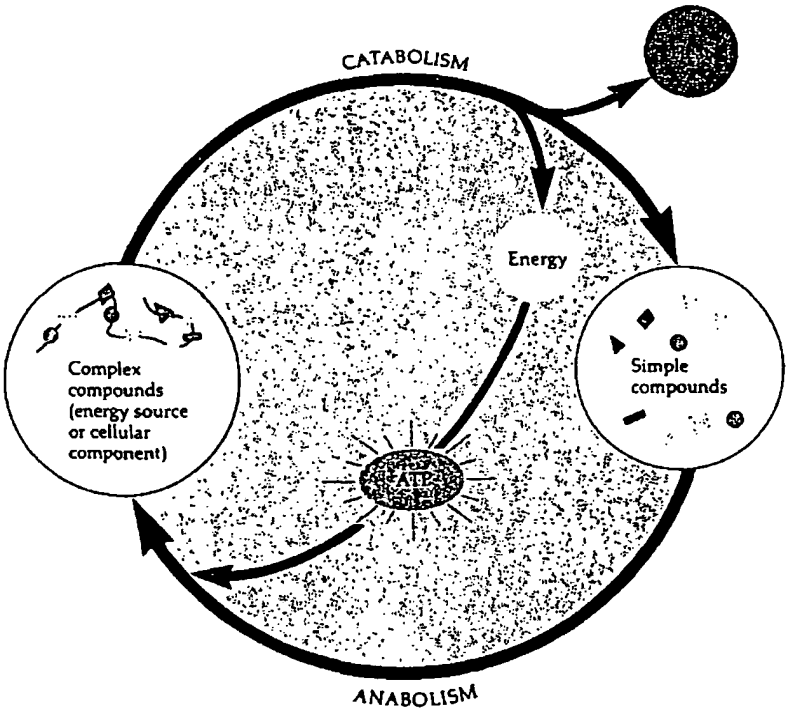


Figure 5–1 Relationship between anabolism and catabolism and the role of ATP. When simple compounds are combined to form complex compounds (anabolism), ATP provides the energy for synthesis. When large compounds are split apart (catabolism), heat energy is given off and some energy is trapped in ATP molecules.

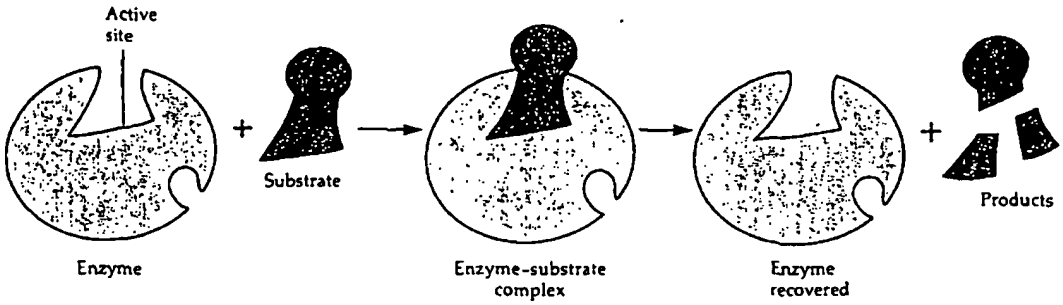


Figure 5–2 Mechanism of enzyme action. The surface of the substrate comes into contact with the active site on the surface of the enzyme to form an enzyme–substrate complex. The substrate is then transformed into products and the enzyme is recovered.

changed by it. Generally large globular proteins, enzymes range in molecular weight from about 10,000 to somewhere in the millions. Of the thousand or more known enzymes, each has a three-dimensional characteristic shape with a specific surface configuration due to its primary, secondary, and tertiary structures (see Figure 2–18).

Mechanism of Enzyme Action

As mentioned in Chapter 2, catalysts lower the *activation energy* required for a chemical reaction. Although scientists do not completely understand how an enzyme does this, the sequence of events is believed to be as follows (Figure 5–2):

1. The surface of the *substrate*—that is, the molecule or molecules that are reactants in the chemical reaction to be catalyzed—contacts a specific region on the surface of the enzyme molecule, called the *active site*.
2. A temporary intermediate compound called an *enzyme–substrate complex* forms.
3. The substrate molecule is transformed (by rearrangement of existing atoms, a breakdown of the substrate molecule, or the combining of several substrate molecules).
4. The transformed substrate molecules, the products of the reaction, move away from the surface of the enzyme molecule.
5. The recovered enzyme, now freed, reacts with other substrate molecules.

Enzyme reaction is characterized by its extreme *specificity* for a particular substrate. For example, a specific enzyme may be capable of hydrolyzing a peptide bond only between two specific amino acids. And other enzymes are capable of hydrolyzing starch, but not cellulose; even though both starch and cellulose are polysaccharides composed of glucose subunits, the orientations of the subunits in the two polysaccharides differ. Enzyme specificity results from the three-dimensional shape of the active site, which fits the substrate somewhat like a lock with its key. In most instances, the substrate is much smaller than the enzyme, and relatively few of the enzyme's amino acids make up the active site.

A given compound can be a substrate for a number of different enzymes that catalyze different reactions. The fate of a given reactant (substrate) depends on the specific enzyme that reacts upon it. For example, glucose-6-phosphate, an important molecule in cell metabolism, may be acted upon by at least four different enzymes, each of which will give a different product.

Enzymes are exceedingly efficient. Under optimum conditions, they can catalyze reactions at rates that are 10^8 to 10^{10} times (up to 10 billion times) more rapid than those of comparable reactions without enzymes. The *turnover number* (number of substrate molecules metabolized per enzyme mol-

Volume I

Todd • Sanford • Davidsohn

CLINICAL DIAGNOSIS *and* MANAGEMENT by LABORATORY METHODS

Sixteenth Edition

JOHN BERNARD HENRY, M.D.

Dean and Professor of Pathology, Georgetown University School of Medicine; Attending Pathologist, University Hospital, Georgetown University, Washington, D.C.; Formerly Professor of Pathology, College of Medicine, and Director of Clinical Pathology, University Hospital, State University of New York, Upstate Medical Center, Syracuse

W. B. SAUNDERS COMPANY
Philadelphia London Toronto

W. B. Saunders Company: West Washington Square
Philadelphia, PA 19105

1 St. Anne's Road
Eastbourne, East Sussex BN 21 3UN, England

1 Goldthorne Avenue
Toronto, Ontario M8Z 5T9, Canada

Library of Congress Cataloging in Publication Data

Todd, James Campbell

Clinical diagnosis by laboratory methods.

At head of title: Todd-Sanford-Davidsohn.

Includes bibliographies.

1. Diagnosis, Laboratory. I. Sanford, Arthur Hawley, 1882- II. Davidsohn, Israel, 1895- III. Henry, John Bernard, 1928- IV. Title. [DNLM: 1. Diagnosis, Laboratory. QY4.3 T633]

RB37.T63 1979 616.07'56 77-92666

ISBN 0-7216-4639-5

Clinical Diagnosis and Management by Laboratory Methods

Volume 1 0-7216-4652-2
Volume 2 0-7216-4653-0
Single Volume 0-7216-4654-9
Set ISBN 0-7216-4639-5

Last digit is the print number: 9 8 7 6 5 4

destruction of the red cells with higher concentrations of the abnormal hemoglobin or selective removal of the abnormal hemoglobin from the cell.

In *heterozygous alpha hemoglobinopathies*, the abnormality in the alpha chain will affect all three hemoglobin types. Therefore, six different hemoglobin types are found—the three normal hemoglobins and the three abnormal forms. Examples are Hb $D_{Baltimore}$, Hb Ann Arbor, and Hb M_{Boston}.

Combinations of abnormalities exist. *Double heterozygotes for two beta chain abnormalities* produce two different abnormal beta chains; therefore, there are two abnormal hemoglobins and no hemoglobin A. An example of this is Hb S-C disease. Double heterozygotes for beta and delta chain abnormalities and for alpha and beta chain abnormalities are rare but have provided important information. The latter will have four major hemoglobin types on electrophoresis: $\alpha_2^A\beta_2^A$; $\alpha_2^X\beta_2^A$; $\alpha_2^A\beta_2^Y$; and $\alpha_2^X\beta_2^Y$.

Double heterozygotes for beta hemoglobinopathy and beta thalassemia are well known. Here, the quantity of abnormal hemoglobin exceeds the normal hemoglobin, in contrast to the heterozygous beta hemoglobinopathies, in which the reverse is true. Examples are Hb S thalassemias and Hb E thalassemia.

Beta hemoglobinopathies

Hemoglobins S, C, D, and E are believed to be polymorphisms because their frequency is greater than can be explained by mutation alone (Lehmann, 1977). They occur in homozygous as well as heterozygous form and involve the beta chain.

Sickle Cell Disease. Homozygous Hb S disease is a serious chronic hemolytic anemia, first manifest in early childhood and often fatal before the age of 30 years. With modern medical care, however, many patients live longer. Hemoglobin S is found almost exclusively in the black population; 0.1 to 0.2 per cent of the blacks born in the United States have sickle cell anemia (Schneider, 1976).

In hemoglobin S the glutamic acid in the sixth position on the beta chain is replaced by valine. This substitution is on the surface of the molecule and changes its charge and, hence, its electrophoretic mobility. Hemoglobin S is freely soluble when fully oxygenated; when oxygen is removed from Hb S, polymerization of the abnormal hemoglobin occurs, forming tactoids (fluid crystals) which are rigid and deform the cell into the shape which gave the cell its name (Fig. 29-7). In homozygous Hb S disease, sickling occurs at physiologic oxygen tensions and the rigidity of the red cells is responsible for the hemolysis as well as for most of the complications. The rigid cells are more vulnerable to trauma and are readily trapped by the reticuloendothelial system, especially the spleen, accounting for the hemolysis. As a result of the hemolysis, severe continued marrow hyperplasia during childhood produces bone changes: expansion of the marrow space, thinning of the cortex, and radial striations seen in the skull on x-ray. Leg ulcers are common.

COMPLICATIONS. In early childhood, bilateral painful swelling of the dorsa of the hands or feet occurs as a result of sickling and capillary stasis; this is known as the *hand-foot syndrome* or sickle cell dactylitis. It lasts about two weeks, is accompanied by changes of periostitis as observed by x-ray, and does not occur after the age of four.

The spleen is central to three complications: A *sequestration crisis* refers to sudden pooling of blood and rapid enlargement of the spleen, resulting in hypovolemic shock. This may occur in early childhood when splenomegaly is present. *Functional asplenia* (Pearson, 1969) consists of inadequate antibody responses under some conditions and an impaired ability of the reticuloendothelial system to clear bacteria and particulate material from the blood, probably due to reticuloendothelial blockade. This may partly explain the increased risk of infection in children with the disease. Salmonella and pneumococcal infections are unusually prevalent in children with sickle cell anemia. *Autosplenectomy* is the result of vaso-occlusive episodes, resulting in progressive infarction, fibrosis, and contraction of the spleen. Though splenomegaly is present in childhood, a small fibrotic remnant is the rule in the adult.

From early childhood, patients cannot produce a concentrated urine, apparently as a result of anoxic damage to the vasa recta in the medullae of the kidneys. Hematuria as a result of papillary necrosis is common.

Vaso-occlusive crises are debilitating episodes of abdominal and bone or joint pain, accompanied by fever, which are probably due to plugging of small blood vessels by masses of sickled cells. Bone necrosis occurs and may be a focus for salmonella osteomyelitis. Aseptic necrosis of the femoral head is occasionally a complication. The various complications as a